

Quantum Finite Automata and Probabilistic Reversible Automata: R-trivial Idempotent Languages ^{*}

Marats Golovkins, Maksim Kravtsev, and Vasilij Kravcevs

Faculty of Computing, University of Latvia, Raiņa bulv. 19, Riga LV-1586, Latvia
 marats AT latnet DOT lv, maksims DOT kravcevs AT lu DOT lv,
 kvasilijs AT gmail DOT com

Abstract. We study the recognition of \mathcal{R}_1 -trivial idempotent (\mathcal{R}_1) languages by various models of "decide-and-halt" quantum finite automata (QFA) and probabilistic reversible automata (DH-PRA). We introduce *bistochastic* QFA (MM-BQFA), a model which generalizes both Nayak's enhanced QFA and DH-PRA. We apply tools from algebraic automata theory and systems of linear inequalities to give a complete characterization of \mathcal{R}_1 languages recognized by all these models. We also find that "forbidden constructions" known so far do not include all of the languages that cannot be recognized by measure-many QFA.

1 Introduction

Measure-many quantum finite automata (MM-QFA) were defined in 1997 [22] and since then, their language class characterization problem remains open. The difficulties arise because the language class is not closed under Boolean operations like union and intersection [5]. Later on, a probabilistic reversible ("decide-and-halt" probabilistic reversible automaton, DH-PRA) and a more general model of quantum finite automata (enhanced quantum finite automaton, EQFA) were defined as well, which remarkably share with MM-QFA the same property of non-closure [16, 26].

Nevertheless, other probabilistic reversible and quantum models of finite automata are known as well ("classical" probabilistic reversible automata, C-PRA, and Latvian quantum finite automata, LQFA), closed under Boolean operations [15, 2]. The language class characterization problem for these models were solved by help of algebraic automata theory [2]. As a matter of fact, the language classes of both models form the same language variety, corresponding to the **EJ** monoid variety.

In [2], it is also stated that MM-QFA recognize any regular language corresponding to the monoid variety **EJ**. Since any syntactic monoid of a unary regular language belongs to **EJ**, the results in [2] imply that MM-QFA recognize any unary regular language. In [9], MM-QFA recognizing unary languages

^{*} Supported by the Latvian Council of Science, grant No. 09.1570 and by the European Social Fund, contract No. 2009/0216/1DP/1.1.1.2.0/09/APIA/VIAA/044.

are studied in detail, the authors give a new proof of this result by explicitly constructing MM-QFA recognizing unary languages.

The results by Brodsky and Pippenger [10] combined with the non-closure property imply that the class of languages recognized by MM-QFA is a proper subclass of the language variety corresponding to the **ER** monoid variety. The same holds for DH-PRA and for EQFA [16, 26]. In the paper, we consider a sub-variety of **ER**, the variety of \mathcal{R} -trivial idempotent monoids **R₁** and determine which are the \mathcal{R} -trivial idempotent languages (\mathcal{R}_1 languages) that are recognizable by DH-PRA, MM-QFA, EQFA and MM-BQFA ("decide-and-halt" models). Since **R₁** shares a lot of the characteristic properties with **ER**, the obtained results may serve as an insight to solve the general problem relevant to **ER**.

The paper is structured as follows. Section 2 gives definitions used throughout the paper. Section 3 describes the algebraic tools - monoids, morphisms and varieties. Section 4 considers completely positive maps. We apply von Neumann-Halperin theorem and the result by Kuperberg to obtain Theorem 4.7, which is essential to prove the limitations of QFA in terms of language recognition. Sections 5, 6, 7, 8, 9 present the main results of the paper:

- (1) Introduction of MM-BQFA, a model which generalizes the earlier "decide-and-halt" automata models (Section 5, Definition 5.1);
- (2) Definition of systems of linear inequalities corresponding to \mathcal{R}_1 languages. Proof that any \mathcal{R}_1 language cannot be recognized by the "decide-and-halt" models, if its system of linear inequalities is not consistent. (Section 6, Definition 6.4, Theorem 6.5);
- (3) The construction of DH-PRA (this presumes also EQFA and MM-BQFA) and MM-QFA for any \mathcal{R}_1 language having a consistent system of inequalities. Consequently, we obtain that all four "decide-and-halt" models recognize exactly the same \mathcal{R}_1 languages. An \mathcal{R}_1 language is recognizable by any of these models if and only if the corresponding system of linear inequalities is consistent. (Sections 7, 8, Theorems 7.2, 8.3, 8.4, 8.5);
- (4) The proof that the "forbidden constructions" known from [5] do not give all of the languages that cannot be recognized by MM-QFA (Section 9, Theorem 9.2).

Among other results, we obtain the language class recognized by MO-BQFA (Theorem 5.2) and give some closure properties of MM-BQFA (Theorems 5.3, 5.4 and Corollaries 6.6, 7.3).

2 Preliminaries

Given an alphabet A , let A^* be the set of words over alphabet A . Given a word \mathbf{x} , let $|\mathbf{x}|$ be the length of \mathbf{x} . Introduce a partial order \leq on A^* , let $\mathbf{x} \leq \mathbf{y}$ if and only if there exists $\mathbf{z} \in A^*$ such that $\mathbf{xz} = \mathbf{y}$.

Let $\mathcal{P}(A)$ be the set of subsets of A , including the empty set \emptyset . Note that there is a natural partial order on $\mathcal{P}(A)$, i.e., the subset order. Given a word

$\mathbf{s} \in A^*$, let $\mathbf{s}\omega$ be the set of letters of the word \mathbf{s} . We say that $\mathbf{u}, \mathbf{v} \in A^*$ are equivalent with respect to ω , $\mathbf{u} \sim_\omega \mathbf{v}$, if $\mathbf{u}\omega = \mathbf{v}\omega$ (that is, \mathbf{u} and \mathbf{v} consist of the same set of letters). Note that \sim_ω is an equivalence relation. The function ω is a morphism; $(\mathbf{uv})\omega = \mathbf{u}\omega \cup \mathbf{v}\omega$. Moreover, ω preserves the order relation since $\mathbf{u} \leq \mathbf{v}$ implies $\mathbf{u}\omega \subseteq \mathbf{v}\omega$.

Let $\mathcal{F}(A)$ be the set of all words over the alphabet A that do not contain any repeated letters. The empty word ε is an element of $\mathcal{F}(A)$. Let τ be a function such that for every $\mathbf{s} \in A^*$, any repeated letters in \mathbf{s} are deleted, leaving only the first occurrence. We say that $\mathbf{u}, \mathbf{v} \in A^*$ are equivalent with respect to τ , $\mathbf{u} \sim_\tau \mathbf{v}$, if $\mathbf{u}\tau = \mathbf{v}\tau$. Note that \sim_τ is an equivalence relation. Introduce a partial order \leq on $\mathcal{F}(A)$, let $\mathbf{v}_1 \leq \mathbf{v}_2$ if and only if there exists $\mathbf{v} \in \mathcal{F}(A)$ such that $\mathbf{v}_1\mathbf{v} = \mathbf{v}_2$.

A deterministic finite automaton \mathcal{A} is a tuple (Q, A, q_0, \cdot) , where Q - a set of states, A - a finite alphabet, q_0 - an initial state and \cdot is a transition function, that is, an everywhere defined function from $Q \times A$ to Q . We say that a state q of the automaton \mathcal{A} *accepts* a word $\mathbf{x} \in A^*$, if the input \mathbf{x} sets \mathcal{A} into the state q . Given an automaton (Q, A, q_0, \cdot) , one may assign to it a set of final states Q_F , a subset of Q . The resulting automaton is denoted by (Q, A, q_0, \cdot, Q_F) .

3 Monoids and Varieties

A general overview on varieties of finite semigroups, monoids as well as operations on them is given in [33]. It can also serve as a source for the definitions of morphisms and word quotients.

Unless specified otherwise, the monoids discussed in this section are assumed to be finite.

An element e of a monoid \mathcal{M} is called an *idempotent*, if $e^2 = e$. It is a well-known fact that for any monoid \mathcal{M} there exists $k > 0$ such that for any element $x \in \mathcal{M}$ x^k is idempotent. Moreover, if x^k and x^l both are idempotents, then $x^k = x^l$. If x is an element of a monoid \mathcal{M} , the unique idempotent of the subsemigroup of \mathcal{M} generated by x is denoted by x^ω . The set of idempotents of the monoid \mathcal{M} is denoted by $E(\mathcal{M})$.

Given a regular language $L \subseteq A^*$, words $\mathbf{u}, \mathbf{v} \in A^*$ are called *syntactically congruent*, $\mathbf{u} \sim_L \mathbf{v}$, if for all $\mathbf{x}, \mathbf{y} \in A^*$ $\mathbf{xuy} \in L$ if and only if $\mathbf{xvy} \in L$. The set of equivalence classes A^*/\sim_L is a monoid, called *syntactic monoid* of L and denoted $\mathcal{M}(L)$. The morphism φ from A^* to A^*/\sim_L is called *syntactic morphism*.

Given a monoid variety \mathbf{V} , the corresponding language variety is denoted by \mathbf{V} . The set of languages over alphabet A recognized by monoids in \mathbf{V} is denoted by $A^*\mathbf{V}$.

3.1 Varieties Definitions

The monoid varieties used in this paper may be defined by some simple identities. For example, a monoid \mathcal{M} belongs to the variety defined by an identity $\llbracket xy = yx \rrbracket$ if and only if for any $x, y \in \mathcal{M}$ $xy = yx$. In this paper, we shall refer to the following monoid varieties:

- (1) $\mathbf{G} = \llbracket x^\omega = 1 \rrbracket$, the variety of groups.
The respective language variety is denoted \mathcal{G} ;
- (2) $\mathbf{J}_1 = \llbracket x^2 = x, xy = yx \rrbracket$, the variety of commutative and idempotent monoids, also known as semilattice monoids.
The respective language variety - \mathcal{J}_1 (semilattice languages);
- (3) $\mathbf{R}_1 = \llbracket xyx = xy \rrbracket$, the variety of \mathcal{R} -trivial idempotent monoids, also known as left regular band monoids. The respective language variety - \mathcal{R}_1 (\mathcal{R} -trivial idempotent languages, or \mathcal{R}_1 languages);
- (4) $\mathbf{ER}_1 = \llbracket x^\omega y^\omega x^\omega = x^\omega y^\omega \rrbracket$, the variety of such monoids \mathcal{M} that $E(\mathcal{M})$ is an \mathcal{R} -trivial idempotent monoid. This variety is equal to $\mathbf{R}_1 * \mathbf{G}$ [17], the variety generated by semidirect products of \mathcal{R} -trivial idempotent monoids by groups. The respective language variety - \mathcal{ER}_1 ;
- (5) $\mathbf{J} = \llbracket x^\omega x = x^\omega, (xy)^\omega = (yx)^\omega \rrbracket = \llbracket (xy)^\omega x = (xy)^\omega, x(yx)^\omega = (yx)^\omega \rrbracket$, the variety of \mathcal{J} -trivial monoids. The respective language variety - \mathcal{J} ;
- (6) $\mathbf{R} = \llbracket (xy)^\omega x = (xy)^\omega \rrbracket$, the variety of \mathcal{R} -trivial monoids.
The respective language variety - \mathcal{R} ;
- (7) $\mathbf{EJ} = \llbracket (x^\omega y^\omega)^\omega = (y^\omega x^\omega)^\omega \rrbracket = \llbracket (x^\omega y^\omega)^\omega x^\omega = (x^\omega y^\omega)^\omega, x^\omega (y^\omega x^\omega)^\omega = (y^\omega x^\omega)^\omega \rrbracket$, the variety of such monoids \mathcal{M} that $E(\mathcal{M})$ generates a \mathcal{J} -trivial monoid. This variety is equal to $\mathbf{J} * \mathbf{G}$, the variety generated by semidirect products of \mathcal{J} -trivial monoids by groups [34].
The respective language variety - \mathcal{EJ} ;
- (8) $\mathbf{ER} = \llbracket (x^\omega y^\omega)^\omega x^\omega = (x^\omega y^\omega)^\omega \rrbracket$, the variety considered in [14]. It is the variety of such monoids \mathcal{M} that $E(\mathcal{M})$ generates an \mathcal{R} -trivial monoid [1, p.132]. This variety is equal to $\mathbf{R} * \mathbf{G}$, the variety generated by semidirect products of \mathcal{R} -trivial monoids by groups [1, p.344].
The respective language variety - \mathcal{ER} .

It is possible to check that $\mathbf{J}_1 \subset \mathbf{J} \subset \mathbf{EJ}$, $\mathbf{R}_1 \subset \mathbf{R} \subset \mathbf{ER}$, $\mathbf{R}_1 \subset \mathbf{ER}_1 \subset \mathbf{ER}$, $\mathbf{J}_1 \subset \mathbf{R}_1$, $\mathbf{J} \subset \mathbf{R}$ and $\mathbf{G} \subset \mathbf{EJ} \subset \mathbf{ER}$.

3.2 Semilattice Languages and Free Semilattices

We need some characterizations for semilattice languages.

Definition 3.1. A free semilattice over an alphabet A is a monoid $(\mathcal{P}(A), \cup)$, where \cup is the ordinary set union.

For any alphabet A , the free semilattice $\mathcal{P}(A)$ satisfies the identities of \mathbf{J}_1 , therefore $\mathcal{P}(A) \in \mathbf{J}_1$.

For the sake of completeness, we give a proof for the following

Proposition 3.2. Given a language $L \in A^* \mathcal{J}_1$, the free semilattice $\mathcal{P}(A)$ is divided by $\mathcal{M}(L)$.

Proof. Let φ be the syntactic morphism from A^* to $\mathcal{M}(L)$. It suffices to prove that $\omega^{-1}\varphi$ is a surjective morphism.

Let $\mathbf{s}_1, \mathbf{s}_2 \in A^*$. Since $\mathcal{M}(L) \in \mathbf{J}_1$, $\mathbf{s}_1 \sim_\omega \mathbf{s}_2$ implies $\mathbf{s}_1\varphi = \mathbf{s}_2\varphi$. Let $p \in \mathcal{P}(A)$. Let $\mathbf{t}_1, \mathbf{t}_2 \in p\omega^{-1}$. Now, since $\mathbf{t}_1 \sim_\omega \mathbf{t}_2$, $\mathbf{t}_1\varphi = \mathbf{t}_2\varphi$. Hence $\omega^{-1}\varphi$ is a function.

Let $p_1, p_2 \in \mathcal{P}(A)$. Let $\mathbf{s}_1 \in (p_1 p_2) \omega^{-1}$ and let $\mathbf{s}_2 \in (p_1 \omega^{-1})(p_2 \omega^{-1})$. The words \mathbf{s}_1 and \mathbf{s}_2 consist of the same set of letters, so $\mathbf{s}_1 \sim_\omega \mathbf{s}_2$. Therefore $\mathbf{s}_1 \varphi = \mathbf{s}_2 \varphi$, hence $((p_1 p_2) \omega^{-1}) \varphi = ((p_1 \omega^{-1})(p_2 \omega^{-1})) \varphi = p_1(\omega^{-1} \varphi) p_2(\omega^{-1} \varphi)$. So $\omega^{-1} \varphi$ is a morphism.

The morphism φ is surjective and ω is everywhere defined, therefore $\omega^{-1} \varphi$ is surjective. \square

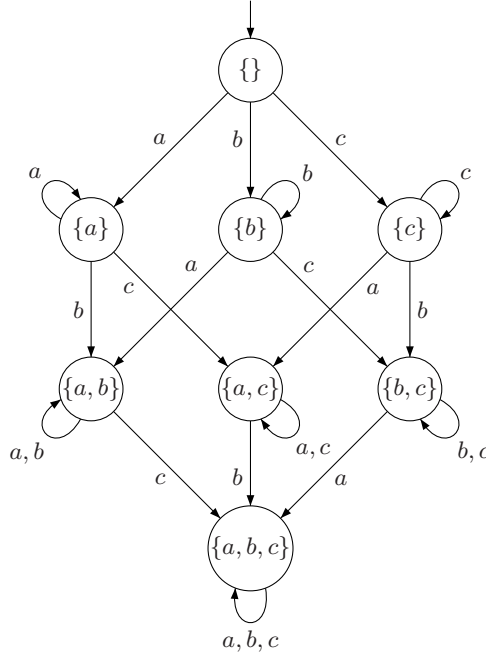


Fig. 1. Free semilattice over $\{a, b, c\}$.

An immediate consequence [33, p.17, Prop. 2.7] is that $\mathcal{P}(A)$ recognizes any language L in $A^* \mathcal{J}_1$. Moreover, L is a disjoint union of some languages $X_1 \omega^{-1}, \dots, X_n \omega^{-1}$, where $X_1, \dots, X_n \in \mathcal{P}(A)$.

Thus, taking into account [33, p.40, Prop. 3.10], the following characterizations have been established:

Theorem 3.3. *Let L be a language over alphabet A . The following conditions are equivalent:*

- (1) *The syntactic monoid of L belongs to the variety \mathbf{J}_1 ;*
- (2) *L is a Boolean combination of languages of the form $A^* a A^*$, where $a \in A$;*
- (3) *L is a Boolean combination of languages of the form B^* , where $B \subseteq A$;*
- (4) *L is a disjoint union of languages of the form $X_1 \omega^{-1}, \dots, X_n \omega^{-1}$, where $X_1, \dots, X_n \in \mathcal{P}(A)$.*

Therefore, in order to specify a particular language $L \in A^* \mathcal{J}_1$, one may identify it by indicating a particular subset of $\mathcal{P}(A)$.

Given a free semilattice $\mathcal{P}(A)$, one may represent it as a deterministic finite automaton $(\mathcal{P}(A), A, \emptyset, \cdot)$, where for every $X \in \mathcal{P}(A)$ and for every $a \in A$,

$X \cdot a = X \cup \{a\}$. By Theorem 3.3 (4), for any semilattice language L over alphabet A , $L\omega$ is a set of final states, such that the automaton recognizes the language.

A free semilattice over $\{a, b, c\}$ represented as a finite automaton is depicted in Figure 1.

The states of $(\mathcal{P}(A), A, \emptyset, \cdot)$ can be separated into several levels, i.e., a state is at level k if it corresponds to an element in $\mathcal{P}(A)$ of cardinality k .

3.3 \mathcal{R}_1 languages and Free Left Regular Bands

We also need some characterizations for \mathcal{R}_1 languages.

Definition 3.4. *A free left regular band over an alphabet A is a monoid $(\mathcal{F}(A), \cdot)$, where $\mathbf{x} \cdot \mathbf{y} = (\mathbf{xy})\tau$, i.e., concatenation followed by the application of τ .*

The function τ is a morphism; for any $\mathbf{u}, \mathbf{v} \in A^*$ $(\mathbf{uv})\tau = \mathbf{u}\tau \cdot \mathbf{v}\tau$. Moreover, τ preserves the order relation since $\mathbf{u} \leq \mathbf{v}$ implies $\mathbf{u}\tau \leq \mathbf{v}\tau$.

For any alphabet A , the free left regular band $\mathcal{F}(A)$ satisfies the identities of \mathbf{R}_1 , therefore $\mathcal{F}(A) \in \mathbf{R}_1$.

Characterizations of \mathcal{R}_1 languages are established in [35]:

Theorem 3.5. *Let L be a language over alphabet A . The following conditions are equivalent:*

- (1) *The syntactic monoid of L belongs to the variety \mathbf{R}_1 ;*
- (2) *L is a Boolean combination of languages of the form B^*aA^* , where $a \in A$ and $B \subseteq A$;*
- (3) *L is a disjoint union of languages of the form*

$$a_1 a_1^* a_2 \{a_1, a_2\}^* a_3 \{a_1, a_2, a_3\}^* \dots a_m \{a_1, a_2, \dots, a_m\}^*,$$

where the a_i 's are distinct letters of A .

Let L be a single language from the disjoint union specified in Theorem 3.5 (3). There exists a single element $\mathbf{x} \in \mathcal{F}(A)$ such that $\mathbf{x}\tau^{-1} = L$, therefore $\mathcal{F}(A)$ recognizes any language in $A^*\mathbf{R}_1$. Hence by [33, p.17, Prop. 2.7], $\mathcal{M}(L)$ divides $\mathcal{F}(A)$.

Therefore, in order to specify a particular language $L \in A^*\mathbf{R}_1$, one may identify it by indicating a particular subset of $\mathcal{F}(A)$. For example, the semilattice language A^*aA^* may also be denoted as $\{\mathbf{a}, \mathbf{ab}, \mathbf{ba}, \mathbf{ac}, \mathbf{ca}, \mathbf{abc}, \mathbf{acb}, \mathbf{bac}, \mathbf{bca}, \mathbf{cab}, \mathbf{cba}\}$.

It is also self-evident that $\mathcal{P}(A)$ is a quotient of $\mathcal{F}(A)$. Indeed, let σ be a restriction of ω to $\mathcal{F}(A)$. The function σ is a surjective morphism from $\mathcal{F}(A)$ to $\mathcal{P}(A)$ which preserves the order relation.

Given a free left regular band $\mathcal{F}(A)$, one may represent it as a deterministic finite automaton $(\mathcal{F}(A), A, \varepsilon, \cdot_{\mathcal{F}(A)})$. By Theorem 3.5 (3), for any \mathcal{R}_1 language L over alphabet A , $L\tau$ is a set of final states, such that the automaton recognizes the language.

A free left regular band over $\{a, b, c\}$ represented as a finite automaton is depicted in Figure 2.

Free left regular bands and free semilattices are key elements to prove that a quantum automaton may recognize a particular \mathcal{R}_1 language if and only if its system of linear inequalities is consistent.

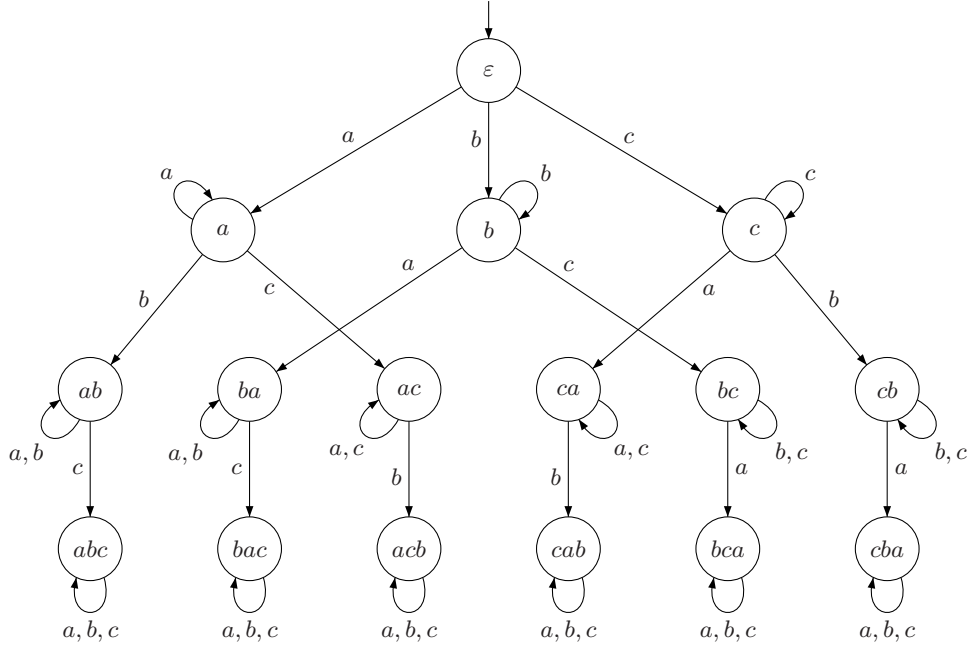


Fig. 2. Free left regular band over $\{a, b, c\}$.

4 Completely Positive Maps

In this section, we establish some facts about completely positive maps with certain properties, i.e., completely positive maps that describe the evolution of bistochastic quantum finite automata, defined in the next section. A comprehensive account on quantum computation can be found in [29].

Following [29], we call a matrix $M \in \mathbb{C}^{n \times n}$ *positive*, if for any vector $X \in \mathbb{C}^n$, X^*MX is real and nonnegative. In literature, positive matrices sometimes are called positive semi-definite. For arbitrary matrices M, N we may write $M \geq N$ if $M - N$ is positive. This defines a partial ordering on $\mathbb{C}^{n \times n}$. Also note that the set of all positive matrices in $\mathbb{C}^{n \times n}$ is an additive monoid. A matrix is positive if and only if it is Hermitian and all of its eigenvalues are nonnegative ([29, Exercises 2.21, 2.24]). A matrix M is positive if and only if exists a matrix $S \in \mathbb{C}^{n \times n}$ such that $M = S^*S$ ([38, Section 6.1]). Let $\text{Tr}(A)$ be a trace of a matrix A . The inner product of two matrices A and B is defined as $\langle A, B \rangle = \text{Tr}(A^*B)$. Consequently, the norm of a matrix A (the Frobenius norm) is defined as $\|A\| = \sqrt{\text{Tr}(A^*A)}$.

Proposition 4.1. *A matrix M is positive if and only if for any positive A $\text{Tr}(MA) \geq 0$.*

Proof. Assume that for any positive A $\text{Tr}(MA) \geq 0$. Take $A = XX^*$, where X is an arbitrary vector. Now $X^*MX = \text{Tr}(MXX^*) \geq 0$, hence for any X $X^*MX \geq 0$. So M is positive.

Assume M is positive. Let A be a positive matrix. So A admits spectral decomposition, $A = \sum_{i=1}^n \lambda_i X_i X_i^*$, where λ_i are nonnegative eigenvalues and X_i - eigenvectors of A . Now $\text{Tr}(MA) = \text{Tr}(M \sum_{i=1}^n \lambda_i X_i X_i^*) = \sum_{i=1}^n \lambda_i \text{Tr}(MX_i X_i^*) = \sum_{i=1}^n \lambda_i X_i^* M X_i \geq 0$. \square

A linear map $\Phi : \mathbb{C}^{n \times n} \rightarrow \mathbb{C}^{m \times m}$ is called *positive*, if for any $n \times n$ positive matrix M $\Phi(M)$ is positive. Any linear map from $\mathbb{C}^{n \times n}$ to $\mathbb{C}^{m \times m}$ may be regarded as a linear operator in $\mathbb{C}^{n^2 \times m^2}$. The norm of a linear map Φ from $\mathbb{C}^{n \times n}$ to $\mathbb{C}^{m \times m}$ is defined as $\sup_{M \in \mathbb{C}^{n \times n}} \frac{\|\Phi(M)\|}{\|M\|}$. A linear map Φ is called a *contraction*, if $\|\Phi\| \leq 1$.

Let I_s be the identity map over $\mathbb{C}^{s \times s}$. Given two linear maps Φ and Ψ , let $\Phi \otimes \Psi$ be the tensor product of those maps. A positive linear map Φ is called *completely positive (CP)*, if for any $s \geq 1$, $\Phi \otimes I_s$ is positive. By Choi's theorem [11], a linear map is completely positive if and only if it admits a *Kraus decomposition*, meaning that there exist matrices $V_1, \dots, V_l \in \mathbb{C}^{m \times n}$, $l \leq nm$, such that for any matrix $M \in \mathbb{C}^{n \times n}$ $\Phi(M) = \sum_{i=1}^l V_i M V_i^*$. So any CP map may be identified by a set of its Kraus operators $\{V_1, \dots, V_l\}$.

A completely positive map Φ is called *trace-preserving*, if for any positive M , $\text{Tr}(\Phi(M)) = \text{Tr}(M)$. A CP map $\Phi = \{V_1, \dots, V_l\}$ from $\mathbb{C}^{n \times n}$ to $\mathbb{C}^{m \times m}$ is trace preserving if and only if $\sum_{i=1}^l V_i^* V_i = I_n$ [29, §8.2.3].

A completely positive map Φ is called *sub-tracial* iff for any positive M we have $\text{Tr}(\Phi(M)) \leq \text{Tr}(M)$.

Theorem 4.2. *A completely positive map $\Phi = \{V_1, \dots, V_l\}$ from $\mathbb{C}^{n \times n}$ to $\mathbb{C}^{m \times m}$ is sub-tracial if and only if $\sum_{i=1}^l V_i^* V_i \leq I_n$.*

Proof. Assume that $\sum_{i=1}^l V_i^* V_i \leq I_n$. So exists a positive matrix P such that

$\sum_{i=1}^l V_i^* V_i + P = I_n$. Moreover, $P = \sum_{i=1}^n \lambda_i X_i X_i^*$, where λ_i are nonnegative eigenvalues and X_i - eigenvectors of P . By adding $m-1$ zero columns to each vector X_i one respectively obtains matrices $W_i \in \mathbb{C}^{n \times m}$ such that $X_i X_i^* = W_i W_i^*$. For each i , $1 \leq i \leq n$, let $V_{l+i} = \sqrt{\lambda_i} W_i^*$. So $\sum_{i=1}^{l+n} V_i^* V_i = I_n$. Hence $\{V_1, \dots, V_l, \dots, V_{l+n}\}$

is a trace-preserving CP map, so for any positive M $\text{Tr}\left(\sum_{i=1}^{l+n} V_i M V_i^*\right) = \text{Tr}(M)$.

The matrix $\sum_{i=1}^n V_{l+i} M V_{l+i}^*$ is positive, therefore $\text{Tr} \left(\sum_{i=1}^n V_{l+i} M V_{l+i}^* \right) \geq 0$. Hence $\text{Tr} \left(\sum_{i=1}^l V_i M V_i^* \right) \leq \text{Tr}(M)$.

Assume that for all positive M $\text{Tr}(\Phi(M)) \leq \text{Tr}(M)$. Since $\text{Tr} \left(\sum_{i=1}^l V_i M V_i^* \right) = \text{Tr} \left(\sum_{i=1}^l V_i^* V_i M \right)$, for all positive M $\text{Tr} \left(\sum_{i=1}^l V_i^* V_i M \right) \leq \text{Tr}(M)$. So for any positive M $\text{Tr} \left(\left(I_n - \sum_{i=1}^l V_i^* V_i \right) M \right) \geq 0$. Now by Proposition 4.1, $I_n - \sum_{i=1}^l V_i^* V_i$ is positive, therefore $\sum_{i=1}^l V_i^* V_i \leq I_n$. \square

A CP map $\Phi = \{V_1, \dots, V_l\}$ from $\mathbb{C}^{n \times n}$ to $\mathbb{C}^{m \times m}$ is called *unital* if $\Phi(I_n) = I_m$, i.e., $\sum_{i=1}^l V_i V_i^* = I_m$. A CP map from $\mathbb{C}^{n \times n}$ to $\mathbb{C}^{m \times m}$ $\Phi = \{V_1, \dots, V_l\}$ is called *sub-unital* if $\Phi(I_n) \leq I_m$, i.e., $\sum_{i=1}^l V_i V_i^* \leq I_m$.

A *composition* of CP maps Φ_0, \dots, Φ_m from $\mathbb{C}^{n \times n}$ to $\mathbb{C}^{n \times n}$ is a CP map $\Phi = \Phi_0 \circ \dots \circ \Phi_m$ such that for any $M \in \mathbb{C}^{n \times n}$ $\Phi(M) = \Phi_0(\Phi_1(\dots(\Phi_m(M))\dots))$.

A CP map $\Phi = \{V_1, \dots, V_l\}$ from $\mathbb{C}^{n \times n}$ to $\mathbb{C}^{n \times n}$ is called *bistochastic*, if it is both trace preserving and unital, i.e., $\sum_{i=1}^l V_i V_i^* = \sum_{i=1}^l V_i^* V_i = I_n$.

Examples of bistochastic CP maps.

- (1) A map defined by unitary matrix U , i.e., a CP map $\Phi(M) = U M U^*$, called *unitary operation*;
- (2) A collection of projection matrices $\{P_i\}$ such that $\sum_{i=1}^l P_i = I$, i.e., a CP map $\Phi(M) = \sum_{i=1}^l P_i M P_i^*$, called *orthogonal measurement*;
- (3) A CP map $\Phi(M) = \sum_{i=1}^l p_i U_i M U_i^*$, where $\sum_{i=1}^l p_i = 1$ and for all i U_i are unitary. Such a map is called *random unitary operation*;
- (4) Any composition of the maps above.

A CP map $\mathbb{C}^{n \times n}$ to $\mathbb{C}^{n \times n}$ is called *sub-bistochastic*¹, if it is both sub-unital and sub-tracial. A composition of two sub-bistochastic CP maps is a sub-bistochastic CP map.

We are interested about some properties of the asymptotic dynamics resulting from iterative application of a CP sub-bistochastic map.

A CP map Φ from $\mathbb{C}^{n \times n}$ to $\mathbb{C}^{n \times n}$ is called *idempotent* if $\Phi \circ \Phi = \Phi$.

¹ Sometimes in quantum physics and quantum computation literature, a CP map is sub-tracial by definition. In such cases, sub-bistochastic CP maps are called sub-unital CP.

Definition 4.3. A CP map Φ from $\mathbb{C}^{n \times n}$ to $\mathbb{C}^{n \times n}$ generates a unique idempotent, denoted Φ^ω , if there exists a sequence of positive integers n_s such that 1) exists the limit $\Phi^\omega = \lim_{s \rightarrow \infty} \Phi^{n_s}$; 2) the CP map Φ^ω is idempotent; 3) for any sequence of positive integers m_s such that the limit $\lim_{s \rightarrow \infty} \Phi^{m_s}$ exists and is idempotent, $\lim_{s \rightarrow \infty} \Phi^{m_s} = \Phi^\omega$.

For example, if Φ is a unitary operation then Φ^ω is the identity map. (Theorem 4.5.)

Note that any CP map from $\mathbb{C}^{n \times n}$ to $\mathbb{C}^{n \times n}$ may be regarded as a linear operator in $\mathbb{C}^{n^2 \times n^2}$. In this sense, the conjugate transpose of $\Phi = \{V_1, \dots, V_l\}$ is $\Phi^* = \{V_1^*, \dots, V_l^*\}$. Kuperberg has provided a sketch of the proof [23] that for any CP sub-bistochastic map Φ from $\mathbb{C}^{n \times n}$ to $\mathbb{C}^{n \times n}$, its idempotent Φ^ω exists and it is a linear projection operator in $\mathbb{C}^{n^2 \times n^2}$. We reconstruct a full proof of this result below. The first step in that direction is the following theorem.

Theorem 4.4. Any CP sub-bistochastic map Φ is a contraction.²

Proof. We need to prove that $\|\Phi\| \leq 1$. Let $\sigma_{\max}(\Phi)$ - the largest singular value of Φ and $\lambda_{\max}(\Phi^* \circ \Phi)$ - the largest eigenvalue of $\Phi^* \circ \Phi$. Note that $\|\Phi\| = \sigma_{\max}(\Phi) = \sqrt{\lambda_{\max}(\Phi^* \circ \Phi)}$. Let M an eigenvector of $\Phi^* \circ \Phi$ corresponding to λ_{\max} . So $\Phi^* \circ \Phi(M) = \lambda_{\max} M$. Suppose M is not Hermitian. Let V_1, \dots, V_l be the Kraus operators corresponding to $\Phi^* \circ \Phi$. So $\Phi^* \circ \Phi(M^*) = \sum_{i=1}^l V_i M^* V_i^* = (\sum_{i=1}^l V_i M V_i^*)^* = (\lambda_{\max} M)^* = \lambda_{\max} M^*$. Hence M^* is an eigenvector corresponding to λ_{\max} as well. Therefore $M + M^*$ is an eigenvector also corresponding to λ_{\max} , and it is Hermitian. So without loss of generality, we may assume that M is Hermitian. Note that $\text{Tr}(\Phi^* \circ \Phi(M)) = \lambda_{\max} \text{Tr}(M)$. On the other hand, since $\Phi^* \circ \Phi$ is sub-bistochastic, $\text{Tr}(\Phi^* \circ \Phi(M)) \leq \text{Tr}(M)$. Hence $\lambda_{\max} \leq 1$. Therefore $\|\Phi\| \leq 1$. \square

Theorem 4.5. Any CP sub-bistochastic map Φ generates a unique idempotent Φ^ω .

Proof. Let σ_{\max} - the largest singular value of Φ and λ - any of its eigenvalues. Due to Browne's theorem [7, Fact 5.11.21 i)], $|\lambda| \leq \sigma_{\max}$. Therefore by Theorem 4.4, $|\lambda| \leq 1$. Let λ_1 an eigenvalue such that $|\lambda_1| = 1$. Let $m(\lambda_1)$ and $g(\lambda_1)$ - the algebraic and geometric multiplicity of λ_1 . It has been proved in [25, Lemmas 2 and 3] that $m(\lambda_1) = g(\lambda_1)$. (The proofs are given for the bistochastic case, but they can be copied for sub-bistochastic case with a sole modification: in the proof of Lemma 2 in [25], replace " $\|\Phi_{\mathcal{A}}\| = 1$ " with " $\|\Phi_{\mathcal{A}}\| \leq 1$ ". Before [25], the same has been proved for random unitary operations in [30].)

The map Φ may be viewed as an $n^2 \times n^2$ matrix, it admits Jordan normal form. So $\Phi = SJS^{-1}$, where J is a Jordan block matrix and S - some non-singular matrix. Consider the Jordan blocks corresponding to any eigenvalue

² The special case dealing with bistochastic maps was proved in [32].

λ_1 such that $|\lambda_1| = 1$. Since $m(\lambda_1) = g(\lambda_1)$, any such Jordan block is one-dimensional. Any other Jordan block B is related to an eigenvalue λ such that $|\lambda| < 1$, so $\lim_{s \rightarrow \infty} B^s = 0$. Consider the diagonal matrix L corresponding to eigenvalues λ_1 such that $|\lambda_1| = 1$. There exists a strictly monotone increasing sequence of positive integers n_s such that $\lim_{s \rightarrow \infty} L^{n_s} = I$. (This is implied by [19, Theorem 201].) Thus $L^\omega = I$. The uniqueness of L^ω comes from the fact that the identity matrix is the only idempotent diagonal matrix with diagonal entries all nonzero. So J generates a unique idempotent; J^ω is a diagonal matrix with zeroes and ones on the diagonal. Therefore Φ generates a unique idempotent as well; $\Phi^\omega = S J^\omega S^{-1}$. \square

If Φ is a CP sub-bistochastic map, then Φ^ω is a CP sub-bistochastic map as well.

Theorem 4.6. *The unique idempotent Φ^ω generated by a CP sub-bistochastic map Φ from $\mathbb{C}^{n \times n}$ to $\mathbb{C}^{n \times n}$ is a projection operator in $\mathbb{C}^{n^2 \times n^2}$.*

Proof. By Theorem 4.4, Φ is a contraction. So Φ^ω is a contraction as well. Therefore, due to Halperin [18, 3.(III)], Φ^ω is a projection. \square

Finally, we are ready to formulate a theorem, which is the main result of this section. As shown further in the paper, this theorem ultimately is the reason why certain models of quantum finite automata cannot recognize all regular languages.

Theorem 4.7. *Let e_1, \dots, e_k be idempotent CP sub-bistochastic maps from $\mathbb{C}^{n \times n}$ to $\mathbb{C}^{n \times n}$. Then for any i , $1 \leq i \leq k$,*

- (1) $\lim_{n \rightarrow \infty} (e_1 \circ \dots \circ e_k)^n = (e_1 \circ \dots \circ e_k)^\omega = (e_{\pi(1)} \circ \dots \circ e_{\pi(k)})^\omega$, where π is a permutation in $\{1, \dots, k\}$;
- (2) $(e_1 \circ \dots \circ e_k)^\omega = e_i \circ (e_1 \circ \dots \circ e_k)^\omega = (e_1 \circ \dots \circ e_k)^\omega \circ e_i$.

Proof. Since e_1, \dots, e_k are projections, by von Neumann-Halperin theorem [18, Theorem 1], $\lim_{n \rightarrow \infty} (e_1 \circ \dots \circ e_k)^n = (e_1 \circ \dots \circ e_k)^\omega = (e_{\pi(1)} \circ \dots \circ e_{\pi(k)})^\omega$. In the same way, $(e_1 \circ \dots \circ e_k)^\omega = (e_i \circ e_1 \circ \dots \circ e_k)^\omega = (e_1 \circ \dots \circ e_k \circ e_i)^\omega$. Note that $e_i \circ (e_i \circ e_1 \circ \dots \circ e_k)^\omega = (e_i \circ e_1 \circ \dots \circ e_k)^\omega$ and $(e_1 \circ \dots \circ e_k \circ e_i)^\omega \circ e_i = (e_1 \circ \dots \circ e_k \circ e_i)^\omega$. Therefore $(e_1 \circ \dots \circ e_k)^\omega = e_i \circ (e_1 \circ \dots \circ e_k)^\omega = (e_1 \circ \dots \circ e_k)^\omega \circ e_i$. \square

Any finite quantum system at a particular moment of time (i.e., its *mixed state*) is described by a *density matrix*. By [29, Theorem 2.5], a matrix is a density matrix if and only if it is positive and its trace is equal to 1.

Informally, an $n \times n$ density matrix describes a quantum system with n states. A completely positive trace-preserving map describes an evolution of a quantum system as allowed by quantum mechanics. It maps a density matrix to a density matrix.

5 Automata Models

An overview of different models of finite automata, relevant to our research, is given in the following table. The definition for bistochastic quantum finite

automata is given below. For the formal definitions of other indicated automata models, the reader is referred to the references given in the table.

As seen further, measure-once (measure-many) bistochastic quantum finite automata is a generalization of any other "classical" ("decide-and-halt", respectively) word acceptance model from Table 1. At the same time BQFA have the same limitations for language recognition as known for other models above. Thus we consider the introduction of yet another model of quantum finite automata justified, because it allows us to prove the limitations of language recognition for all the models within single framework. Therefore the proof of the new limitations for MM-BQFA in Section 6, which are expressed in terms of linear inequalities, implies the same for any other "decide-and-halt" word acceptance model in the table.

Table 1. Automata Models

	"Classical" word acceptance	"Decide-and-halt" word acceptance
Deterministic Reversible Automata	Group Automata (GA) [20, 37]	Reversible Finite Automata (RFA) [3, 17]
Quantum Finite Automata with Pure States	Measure-Once Quantum Finite Automata (MO-QFA) [27, 10]	Measure-Many Quantum Finite Automata (MM-QFA) [22, 10, 5, 2]
Probabilistic Reversible Automata	"Classical" Probabilistic Reversible Automata (C-PRA) [15, 2]	"Decide-and-halt" Probabilistic Reversible Automata (DH-PRA) [15, 16]
Quantum Finite Automata with Mixed States	Latvian Quantum Finite Automata (LQFA) [2]	Enhanced Quantum Finite Automata (EQFA) [28, 26]
	Measure-Once Bistochastic Quantum Finite Automata (MO-BQFA)	Measure-Many Bistochastic Quantum Finite Automata (MM-BQFA)

Definition 5.1. A bistochastic quantum finite automaton (BQFA) is a tuple $(Q, A \cup \{\#, \$\}, q_0, \{\Phi_a\})$, where Q is a finite set of states, A - a finite input alphabet, $\#, \$ \notin A$ - initial and final end-markers, q_0 - an initial state and for each $a \in A \cup \{\#, \$\}$ Φ_a is a CP bistochastic transition map from $\mathbb{C}^{|Q| \times |Q|}$ to $\mathbb{C}^{|Q| \times |Q|}$.

Regardless of which word acceptance model is used, each input word is enclosed into end-markers $\#, \$$. At any step, the mixed state of a BQFA may be described by a density matrix ρ . The computation starts in the state $|q_0\rangle\langle q_0|$.

Operation of a measure-once BQFA and word acceptance. On input letter $a \in A$, ρ is transformed into $\Phi_a(\rho)$. The set of states Q is partitioned into two disjoint subsets Q_{acc} and Q_{rej} . After reading the final end-marker $\$,$ a measurement $\{P_{acc}, P_{rej}\}$ is applied to ρ , where $P_{acc} = \sum_{q \in Q_{acc}} |q\rangle\langle q|$ and $P_{rej} = \sum_{q \in Q_{rej}} |q\rangle\langle q|$.

The respective input word is accepted with probability $\text{Tr}(P_{acc}\rho P_{acc})$ and rejected with probability $\text{Tr}(P_{rej}\rho P_{rej})$. For any word $\mathbf{a} = a_1 \dots a_k$, define $\Phi_{\mathbf{a}} = \Phi_{a_k} \circ \dots \circ \Phi_{a_1}$.

Operation of a measure-many BQFA and word acceptance. The set of states Q is partitioned into three disjoint subsets Q_{non} , Q_{acc} and Q_{rej} - non-halting, accepting and rejecting states, respectively. It is assumed that $q_0 \in Q_{non}$. On input letter $a \in A$, ρ is transformed into $\rho' = \Phi_a(\rho)$. After that, a measurement $\{P_{non}, P_{acc}, P_{rej}\}$ is applied to ρ' , where for each $i \in \{non, acc, rej\}$ $P_i = \sum_{q \in Q_i} |q\rangle\langle q|$. The respective input word is accepted (rejected) with probability $\text{Tr}(P_{acc}\rho'P_{acc})$ ($\text{Tr}(P_{rej}\rho'P_{rej})$, respectively). If the input word is accepted or rejected, the computation is halted. Otherwise, with probability $\text{Tr}(P_{non}\rho'P_{non})$, the computation continues from the mixed state $P_{non}\rho'P_{non} / \text{Tr}(P_{non}\rho'P_{non})$. To ensure that any input word is always either accepted or rejected, it is required for $\Phi_{\$}$ that for any ρ such that $\text{Tr}(P_{non}\rho P_{non}) = 1$, $\text{Tr}(P_{non}\Phi_{\$}(\rho)P_{non}) = 0$.

To describe the probability distribution $S_{\#u}$ of a MM-BQFA \mathcal{A} after reading some prefix $\#u$, it is convenient to use density matrices ρ scaled by p , $0 \leq p \leq 1$. So the probability distribution $S_{\#u}$ of \mathcal{A} is a triple (ρ, p_{acc}, p_{rej}) , where $\text{Tr}(\rho) + p_{acc} + p_{rej} = 1$, $\rho / \text{Tr}(\rho)$ is the current mixed state and p_{acc}, p_{rej} are respectively the probabilities that \mathcal{A} has accepted or rejected the input. So the scaled density matrix ρ may be called a *scaled mixed state*. For any $a \in A \cup \{\#, \$\}$, let $\Psi_a(\rho) = P_{non}\Phi_a(\rho)P_{non}$. After reading the next input letter a , the probability distribution is $S_{\#ua} = (\Psi_a(\rho), p_{acc} + \text{Tr}(P_{acc}\Phi_a(\rho)P_{acc}), p_{rej} + \text{Tr}(P_{rej}\Phi_a(\rho)P_{rej}))$. For any word $\mathbf{a} = a_1 \dots a_k$, define $\Psi_{\mathbf{a}} = \Psi_{a_k} \circ \dots \circ \Psi_{a_1}$. Hence $\rho = \Psi_{\#u}(|q_0\rangle\langle q_0|)$. Note that $\Psi_{\mathbf{a}}$ is a CP sub-bistochastic map.

Language recognition is defined in a way equivalent to Rabin's [36]. Suppose that an automaton \mathcal{A} corresponds to one of the probabilistic or quantum models from the table above. By $p_{\mathbf{x}, \mathcal{A}}$ (or $p_{\mathbf{x}}$, if no ambiguity arises) we denote the probability that an input \mathbf{x} is accepted by the automaton \mathcal{A} . Furthermore, we denote $P_L = \{p_{\mathbf{x}, \mathcal{A}} \mid \mathbf{x} \in L\}$, $\overline{P_L} = \{p_{\mathbf{x}, \mathcal{A}} \mid \mathbf{x} \notin L\}$, $p_1 = \sup \overline{P_L}$, $p_2 = \inf P_L$. It is said that an automaton \mathcal{A} recognizes a language L with interval (p_1, p_2) , if $p_1 \leq p_2$ and $P_L \cap \overline{P_L} = \emptyset$. It is said that an automaton \mathcal{A} recognizes a language L with bounded error and interval (p_1, p_2) , if $p_1 < p_2$. We consider only bounded error language recognition. An automaton is said to recognize a language with probability p if the automaton recognizes the language with interval $(1 - p, p)$. It is said that a language is recognized by some class of automata with probability $1 - \epsilon$, if for every $\epsilon > 0$ there exists an automaton in the class which recognizes the language with interval $(\epsilon_1, 1 - \epsilon_2)$, where $\epsilon_1, \epsilon_2 \leq \epsilon$. A language L is recognizable with interval (p_1, p_2) iff it is recognizable with some probability p (see, for example, [15]).

BQFA as a generalization of other models. Since unitary operations and orthogonal measurements are bistochastic operations, MO-BQFA is a generalization of LQFA and MM-BQFA is a generalization of EQFA. Also one can see that MO-BQFA and MM-BQFA are generalizations of C-PRA and DH-PRA, respectively. A probability distribution vector $P = \sum_i p_i |q_i\rangle$ of a PRA corresponds to the mixed state $\rho = \sum_i p_i |q_i\rangle\langle q_i|$ of a BQFA. Any transition matrix B of a PRA is doubly stochastic. By the Birkhoff theorem [38, Theorem 4.21], any doubly stochastic matrix is a convex combination of some permutation matrices.

Thus $B = \sum_s p_s T_s$, where p_s are nonnegative numbers with sum equal to 1 and T_s - permutation matrices. So the CP bistochastic map corresponding to the transition matrix B is $\Phi(\rho) = \sum_s p_s T_s \rho T_s^*$, which is a random unitary operation.

Indeed, one may check that $\Phi(\rho)$ is a diagonal matrix such that $(\Phi(\rho))_{ii} = (BP)_i$.

On the other hand, BQFA are a special case of one-way general QFA (also called quantum automata with open time evolution), which admit any CP trace-preserving transition maps. One-way general QFA recognize with bounded error exactly the regular languages [21, 24], this fact was also mentioned in [6, Introduction]. Similar models of quantum automata which recognize any regular language have been proposed in [31, 12, 8, 13]. So the recognition power of BQFA is also limited to regular languages only.

Comparison of the language classes. Having a certain class of automata \mathcal{A} , let us denote by $\mathcal{L}(\mathcal{A})$ the respective class of languages. Thus $\mathcal{L}(\text{GA}) = \mathcal{L}(\text{MO-QFA}) = \mathcal{G}$, $\mathcal{L}(\text{C-PRA}) = \mathcal{L}(\text{LQFA}) = \mathcal{L}(\text{MO-BQFA}) = \mathcal{EJ}$, $\mathcal{G} \subsetneq \mathcal{L}(\text{RFA}) \subsetneq \mathcal{ER}_1$, $\mathcal{EJ} \subsetneq \mathcal{L}(\text{MM-QFA}) \stackrel{?}{=} \mathcal{L}(\text{DH-PRA}) \stackrel{?}{=} \mathcal{L}(\text{EQFA}) \stackrel{?}{=} \mathcal{L}(\text{MM-BQFA}) \subsetneq \mathcal{ER}$. Relations concerning BQFA are proved below. All the other relations are known from the references given in Table 1.

Theorem 5.2. $\mathcal{L}(\text{MO-BQFA}) = \mathcal{EJ}$.

Proof. Since $\mathcal{L}(\text{LQFA}) = \mathcal{EJ}$ [2] and MO-BQFA is a generalization of LQFA, $\mathcal{EJ} \subseteq \mathcal{L}(\text{MO-BQFA})$. It remains to prove that $\mathcal{L}(\text{MO-BQFA}) \subseteq \mathcal{EJ}$.

Suppose that a MO-BQFA \mathcal{A} recognizes a language L over alphabet A , such that $L \notin A^* \mathcal{EJ}$. Let $\mathcal{M} = \mathcal{M}(L)$ - the syntactic monoid of L and φ - the syntactic morphism from A^* to \mathcal{M} . By assumption, there exist $x, y \in \mathcal{M}$ such that $(x^\omega y^\omega)^\omega \neq (y^\omega x^\omega)^\omega$. There exists a positive integer k such that for all z in \mathcal{M} $z^k = z^\omega$, therefore $(x^k y^k)^k \neq (y^k x^k)^k$. Let $\mathbf{a} \in x^k \varphi^{-1}$ and $\mathbf{b} \in y^k \varphi^{-1}$. Consider the CP bistochastic transition maps $\Phi_{\mathbf{a}}$ and $\Phi_{\mathbf{b}}$ of \mathcal{A} . Theorem 4.7 implies that there exists a sequence of positive integers s_n such that $\lim_{n \rightarrow \infty} (\Phi_{\mathbf{a}}^{s_n} \circ \Phi_{\mathbf{b}}^{s_n})^n = \lim_{n \rightarrow \infty} (\Phi_{\mathbf{b}}^{s_n} \circ \Phi_{\mathbf{a}}^{s_n})^n = (\Phi_{\mathbf{a}}^\omega \circ \Phi_{\mathbf{b}}^\omega)^\omega = (\Phi_{\mathbf{b}}^\omega \circ \Phi_{\mathbf{a}}^\omega)^\omega$. Note that $\|\Phi_{\mathbf{u}}\| = \|\Phi_{\mathbf{v}}\| = 1$. Therefore for any $\epsilon > 0$ there exists $n > 0$ such that for any $\mathbf{u}, \mathbf{v} \in A^*$ $|p_{\mathbf{u}(\mathbf{a}^{s_n} \mathbf{b}^{s_n})^n \mathbf{v}} - p_{\mathbf{u}(\mathbf{b}^{s_n} \mathbf{a}^{s_n})^n \mathbf{v}}| < \epsilon$. So there exists n such that $(\mathbf{a}^{s_n} \mathbf{b}^{s_n})^n \sim_L (\mathbf{b}^{s_n} \mathbf{a}^{s_n})^n$. Hence $(\mathbf{a}^{s_n} \mathbf{b}^{s_n})^n \varphi = (\mathbf{b}^{s_n} \mathbf{a}^{s_n})^n \varphi = (x^k y^k)^n = (y^k x^k)^n$. The latter implies $(x^k y^k)^k = (y^k x^k)^k$. This is a contradiction. \square

The next theorem is equivalent to the statement that MM-BQFA can't recognize any language that does not satisfy the partial order condition from [10].

Theorem 5.3. $\mathcal{L}(\text{MM-BQFA}) \subseteq \mathcal{ER}$.

Proof. Suppose that a MM-BQFA \mathcal{A} recognizes a language L over alphabet A , such that $L \notin A^* \mathcal{ER}$. Let $\mathcal{M} = \mathcal{M}(L)$ - the syntactic monoid of L and φ - the syntactic morphism from A^* to \mathcal{M} . By assumption, there exist $x, y \in \mathcal{M}$ such that $(x^\omega y^\omega)^\omega x^\omega \neq (x^\omega y^\omega)^\omega$. There exists a positive integer k such that for all z in \mathcal{M} $z^k = z^\omega$, therefore $(x^k y^k)^k x^k \neq (x^k y^k)^k$. Let $\mathbf{a} \in x^k \varphi^{-1}$ and $\mathbf{b} \in y^k \varphi^{-1}$. Consider the CP sub-bistochastic maps $\Psi_{\mathbf{a}}$ and $\Psi_{\mathbf{b}}$ of \mathcal{A} . Theorem 4.7 implies

that there exists a sequence of positive integers s_n such that $\lim_{n \rightarrow \infty} \Psi_{\mathbf{a}}^{s_n} \circ (\Psi_{\mathbf{b}}^{s_n} \circ \Phi_{\mathbf{a}}^{s_n})^n = \lim_{n \rightarrow \infty} (\Psi_{\mathbf{b}}^{s_n} \circ \Psi_{\mathbf{a}}^{s_n})^n = (\Phi_{\mathbf{b}}^\omega \circ \Phi_{\mathbf{a}}^\omega)^\omega$. Let $\mathbf{u}, \mathbf{v} \in A^*$. Let $\mathbf{a}(n) = \mathbf{a}^{s_n}$ and $\mathbf{b}(n) = \mathbf{b}^{s_n}$. Let $\mathbf{w}(n) = \# \mathbf{u}(\mathbf{a}(n)\mathbf{b}(n))^n$. After reading the word $\mathbf{w}(n)$ the probability distribution is $S_{\mathbf{w}(n)} = (\Psi_{\mathbf{w}(n)}(|q_0\rangle\langle q_0|), p_{acc}, p_{rej})$. If $\mathbf{a}(n)$ is read afterwards, $S_{\mathbf{w}(n)\mathbf{a}(n)} = (\Psi_{\mathbf{w}(n)\mathbf{a}(n)}(|q_0\rangle\langle q_0|), p'_{acc}, p'_{rej})$, where $p'_{acc} \geq p_{acc}$ and $p'_{rej} \geq p_{rej}$. Note that $\lim_{n \rightarrow \infty} \Psi_{\mathbf{w}(n)} = \lim_{n \rightarrow \infty} \Psi_{\mathbf{w}(n)\mathbf{a}(n)}$ and $\|\Psi_{\mathbf{u}}\| \leq 1, \|\Psi_{\mathbf{v}}\| \leq 1$. Therefore for any $\epsilon > 0$ there exists $n > 0$ such that for any \mathbf{u} $\|\Psi_{\mathbf{w}(n)}(|q_0\rangle\langle q_0|) - \Psi_{\mathbf{w}(n)\mathbf{a}(n)}(|q_0\rangle\langle q_0|)\| < \epsilon$. The latter in turn implies $p'_{acc} - p_{acc} < \epsilon$ and $p'_{rej} - p_{rej} < \epsilon$. So there exists n such that for any \mathbf{u}, \mathbf{v} $\mathbf{u}(\mathbf{a}^{s_n}\mathbf{b}^{s_n})^n\mathbf{v} \in L$ if and only if $\mathbf{u}(\mathbf{a}^{s_n}\mathbf{b}^{s_n})^n\mathbf{a}^{s_n}\mathbf{v} \in L$. Hence $(\mathbf{a}^{s_n}\mathbf{b}^{s_n})^n\varphi = (\mathbf{a}^{s_n}\mathbf{b}^{s_n})^n\mathbf{a}^{s_n}\varphi = (x^ky^k)^n = (x^ky^k)^nx^k$. The latter implies $(x^ky^k)^k = (x^ky^k)^kx^k$. This is a contradiction. \square

The relation $\mathcal{L}(\text{MM-BQFA}) \subsetneq \mathcal{ER}$ is demonstrated in Section 6 (Corollary 6.6).

$\mathcal{L}(\text{MM-BQFA})$ shares a lot of properties with the language classes of other "decide-and-halt" word acceptance models, like closure under complement and inverse homomorphisms. In Section 9 it is noted that MM-BQFA does not recognize any of the languages corresponding to "forbidden constructions" from [5, Theorem 4.3]. Similarly as other "decide-and-halt" models, $\mathcal{L}(\text{MM-BQFA})$ is not closed under union and intersection.

Theorem 5.4. *The class $\mathcal{L}(\text{MM-BQFA})$ is closed under complement, inverse free monoid morphisms, and word quotient.*

Proof. The proof goes along the same lines as in [10, Theorem 4.1], where the same was proved for $\mathcal{L}(\text{MM-QFA})$. Closure under complement follows from the fact that we can exchange the accepting and rejecting states of the MM-BQFA. Closure under inverse free monoid morphisms is proved in the same way as in [10], it is implied by the deferred measurement principle [29, p.186]. Closure under word quotient is implied by closure under inverse free monoid morphisms and the presence of both end-markers. \square

Non-closure under union and intersection is demonstrated in Section 7 (Corollary 7.3).

6 Linear Inequalities

In this section, we derive a system of linear inequalities that an \mathcal{R}_1 language recognized by a MM-BQFA must satisfy. Let \mathcal{S} be a MM-BQFA over alphabet A . Let $\{\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_R\} = \mathcal{F}(A)$. Assume $\mathbf{v}_0 = \varepsilon$. For any $\mathbf{u} \in A^*$, let $\Psi(\mathbf{u}) = \Psi_{\# \mathbf{u}}$. Recall τ is the natural morphism from A^* to $\mathcal{F}(A)$ (see Section 2 and subsection 3.3). First, we prove that there exist words $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_R \in A^*$, for each i $\mathbf{u}_i\tau = \mathbf{v}_i$, such that the automaton \mathcal{S} has essentially the same scaled density matrices for the words consisting of the same letters:

Proposition 6.1. *For every $\epsilon > 0$ there exists an everywhere defined injective function θ from $\mathcal{F}(A)$ to A^* such that for all $\mathbf{v}, \mathbf{v}' \in \mathcal{F}(A)$*

- (1) $\mathbf{v}\theta\tau = \mathbf{v}$;
- (2) $\mathbf{v} \leq \mathbf{v}'$ if and only if $\mathbf{v}\theta \leq \mathbf{v}'\theta$;
- (3) if $\mathbf{v} \sim_\omega \mathbf{v}'$, then $\|\Psi(\mathbf{v}\theta) - \Psi(\mathbf{v}'\theta)\| < \epsilon$.

Proof. Let m_l ($l = 1, 2, \dots$) be a sequence of positive integers such that for all letters $a \in A$ $\lim_{l \rightarrow \infty} \Psi(a^{m_l}) = \Psi_a^\omega$ (existence is implied by Theorem 4.5 and [19, Theorem 201]).

Let μ be a function that assigns to any word in A^* the same word (of the same length) with letters sorted in alphabetical order. Let \varkappa_i , $i \in \mathbb{N}$, a morphism from A^* to A^* such that for any $a \in A$ $a\varkappa_i = a^i$.

Let $\xi = \xi_l$ be an everywhere defined function from $\mathcal{F}(A)$ to A^* , such that $\varepsilon\xi = \varepsilon$ and for all $\mathbf{v} \in \mathcal{F}(A)$, if $|\mathbf{v}| = 1$ then $\mathbf{v}\xi = \mathbf{v}^{m_l}$ and otherwise, if $|\mathbf{v}| \geq 2$ then $\mathbf{v}\xi = (\mathbf{v}\mu\varkappa_{m_l})^l$.

For any \mathbf{v} in $\mathcal{F}(A)$, where $\mathbf{v} = a_1 \dots a_k$ (a_i are distinct letters of A), define a finite sequence of prefixes, denoted $\mathbf{v}[i]$, where $\mathbf{v}[0] = \varepsilon$ and for all i , $1 \leq i \leq k$, $\mathbf{v}[i] = a_1 \dots a_i$.

Let us define the function $\theta = \theta_l$ by induction as follows. Let $\mathbf{v}[0]\theta = \varepsilon\theta = \varepsilon$ and for all i , $1 \leq i \leq k$, let $\mathbf{v}[i]\theta = (\mathbf{v}[i-1]\theta)(\mathbf{v}[i]\xi)$.

So $\mathbf{v}\theta = \mathbf{v}[k]\theta = \mathbf{v}[1]\xi \dots \mathbf{v}[k]\xi = a_1^{m_l} ((a_1^{m_l} a_2^{m_l})\mu)^l \dots ((a_1^{m_l} a_2^{m_l} \dots a_k^{m_l})\mu)^l$. By construction, $(\mathbf{v}\theta)\tau = a_1 a_2 \dots a_k = \mathbf{v}$.

Consider $\mathbf{v}, \mathbf{v}' \in \mathcal{F}(A)$. Since τ preserves order, $\mathbf{v}\theta \leq \mathbf{v}'\theta$ implies $\mathbf{v} \leq \mathbf{v}'$. Suppose $\mathbf{v} \leq \mathbf{v}'$. By construction, $\mathbf{v}\theta \leq \mathbf{v}'\theta$.

Suppose $\mathbf{v} \sim_\omega \mathbf{v}'$. If $|\mathbf{v}| = |\mathbf{v}'| \leq 1$, the condition (3) of the proposition is satisfied. Hence assume $|\mathbf{v}| = |\mathbf{v}'| \geq 2$. Theorem 4.7 implies that $\lim_{l \rightarrow \infty} \Psi(\mathbf{v}\theta) = \lim_{l \rightarrow \infty} \Psi(\mathbf{v}\xi)$ and $\lim_{l \rightarrow \infty} \Psi(\mathbf{v}'\theta) = \lim_{l \rightarrow \infty} \Psi(\mathbf{v}'\xi)$. Since \mathbf{v} is a permutation of \mathbf{v}' , $\lim_{l \rightarrow \infty} \Psi(\mathbf{v}\xi) = \lim_{l \rightarrow \infty} \Psi(\mathbf{v}'\xi)$. Hence $\lim_{l \rightarrow \infty} \Psi(\mathbf{v}\theta) = \lim_{l \rightarrow \infty} \Psi(\mathbf{v}'\theta)$. Take $\epsilon > 0$. The last equality implies that there exists l such that $\|\Psi(\mathbf{v}\theta) - \Psi(\mathbf{v}'\theta)\| < \epsilon$. Since the monoid $\mathcal{F}(A)$ is finite, there exists l which satisfies the last inequality for any two words \mathbf{v} and \mathbf{v}' in $\mathcal{F}(A)$ such that $\mathbf{v} \sim_\omega \mathbf{v}'$.

Proposition is proved. \square

We are ready to derive the linear inequalities that must be satisfied by \mathcal{S} , if it recognizes an \mathcal{R}_1 language L over alphabet A .

Consider $\mathbf{x} \in A^*$. Suppose $\mathbf{x}\tau = \mathbf{v} = a_1 a_2 \dots a_{|\mathbf{v}|}$. By construction used in the proof of Proposition 6.1, $\mathbf{v}\theta = (\mathbf{v}[1]\xi)(\mathbf{v}[2]\xi) \dots (\mathbf{v}\xi)$. Now $\mathbf{x} \sim_\tau \mathbf{v} \sim_\tau \mathbf{v}\theta$. Therefore $\mathbf{x} \in L$ if and only if $\mathbf{v} \in L$, and if and only if $\mathbf{v}\theta \in L$.

Let us observe how \mathcal{S} processes the input word $\mathbf{v}\theta = (\mathbf{v}[1]\xi)(\mathbf{v}[2]\xi) \dots (\mathbf{v}\xi)$. By the definition of MM-BQFA, any input word is enclosed by end-markers $\#$ and $\$$. Let r_0 be the probability that \mathcal{S} has accepted the input (and halted) after reading the initial end-marker $\#$. For $1 \leq i \leq |\mathbf{v}|$, let $r_{\mathbf{v}[i]}$ be the probability that \mathcal{S} is in a mixed state before reading the first letter of $\mathbf{v}[i]\xi$ and has accepted the input (and halted) after reading $\mathbf{v}[i]\xi$, including the possibility of halting while reading it. Let $g_{\mathbf{v}}$ be the probability that \mathcal{S} is in a mixed state after reading $\mathbf{v}\xi$ and has accepted the input after reading the final end-marker $\$$. It follows that \mathcal{S} accepts $\mathbf{v}\theta$ with probability $p_{\mathbf{v}\theta} = r_0 + r_{\mathbf{v}[1]} + r_{\mathbf{v}[2]} + \dots + r_{\mathbf{v}} + g_{\mathbf{v}}$. Note that

the values $r_{\mathbf{v}[i]}$ and $g_{\mathbf{v}}$ depend on the chosen function θ , which itself depends on the parameter l .

We aim to prove that for any \mathcal{R}_1 language L over alphabet A it is possible to define a linear system of inequalities \mathfrak{L} such that the system is consistent if and only if L can be recognized by MM-BQFA. First of all, it is necessary to define the system itself.

The probabilities $r_{\mathbf{v}[i]}$ can be regarded as symbolic variables (let's call them s-variables) in the formal expression $\hat{p}_{\mathbf{v}\theta} = \hat{r}_0 + \hat{r}_{\mathbf{v}[1]} + \hat{r}_{\mathbf{v}[2]} + \dots + \hat{r}_{\mathbf{v}} + \hat{g}_{\mathbf{v}}$.

Definition 6.2. Two s-variables $\hat{r}_{\mathbf{v}[i]}$ and $\hat{r}_{\mathbf{v}'[j]}$, $1 \leq i \leq |\mathbf{v}|$, $1 \leq j \leq |\mathbf{v}'|$, are called equivalent, $\hat{r}_{\mathbf{v}[i]} \sim \hat{r}_{\mathbf{v}'[j]}$, if $\mathbf{v}[i-1] \sim_{\omega} \mathbf{v}'[j-1]$ and $\mathbf{v}[i] \sim_{\omega} \mathbf{v}'[j]$. Two s-variables $\hat{g}_{\mathbf{v}}$ and $\hat{g}_{\mathbf{v}'}$ are called equivalent, $\hat{g}_{\mathbf{v}} \sim \hat{g}_{\mathbf{v}'}$, if $\mathbf{v} \sim_{\omega} \mathbf{v}'$.

The s-variable \hat{r}_0 is defined to be the only element of the equivalence class $[\hat{r}_0]$. The relation \sim is an equivalence relation in the two sets $\{\hat{r}_{\mathbf{v}[i]} \mid \mathbf{v} \in \mathcal{F}(A) \text{ and } 1 \leq i \leq |\mathbf{v}|\}$ and $\{\hat{g}_{\mathbf{v}} \mid \mathbf{v} \in \mathcal{F}(A)\}$.

If two s-variables $\hat{r}_{\mathbf{v}[i]}$ and $\hat{r}_{\mathbf{v}'[j]}$ are equivalent then $i = j$. Moreover, let's formulate the following

Proposition 6.3. For any $\epsilon > 0$ there exists a function θ from Proposition 6.1 such that for any $\mathbf{v}, \mathbf{v}' \in \mathcal{F}(A)$ and any prefixes $\mathbf{v}[i]$, $\mathbf{v}'[i]$

- (1) if $\hat{r}_{\mathbf{v}[i]} \sim \hat{r}_{\mathbf{v}'[i]}$ then $|r_{\mathbf{v}[i]} - r_{\mathbf{v}'[i]}| < \epsilon$;
- (2) if $\hat{g}_{\mathbf{v}} \sim \hat{g}_{\mathbf{v}'}$ then $|g_{\mathbf{v}} - g_{\mathbf{v}'}| < \epsilon$.

Proof. Suppose $\mathbf{v}[i-1] \sim_{\omega} \mathbf{v}'[i-1]$ and $\mathbf{v}[i] \sim_{\omega} \mathbf{v}'[i]$. In that case, $\mathbf{v}[i]\xi = \mathbf{v}'[i]\xi$. Proposition 6.1 implies that for any ϵ' $\|\Psi(\mathbf{v}[i-1]\theta) - \Psi(\mathbf{v}'[i-1]\theta)\| < \epsilon'$. Hence after reading $\mathbf{v}[i-1]\theta$ or $\mathbf{v}'[i-1]\theta$ the automaton \mathcal{S} comes to essentially the same scaled mixed state. Within a particular step, the probability of accepting the input (and halting) in the future depends only on the current mixed state and the remaining part of the input word. So reading afterwards the word $\mathbf{v}[i]\xi$, which is equal to $\mathbf{v}'[i]\xi$, implies that for any ϵ $|r_{\mathbf{v}[i]} - r_{\mathbf{v}'[i]}| < \epsilon$.

Suppose $|\mathbf{v}| = |\mathbf{v}'|$ and $\mathbf{v} \sim_{\omega} \mathbf{v}'$. Again, after reading the both words $\mathbf{v}\theta$ and $\mathbf{v}'\theta$ the automaton \mathcal{S} is in essentially the same scaled mixed state. So reading the final end-marker yields that for any ϵ $|g_{\mathbf{v}} - g_{\mathbf{v}'}| < \epsilon$. \square

Recall that $\mathcal{F}(A)$ can be viewed as an automaton that recognizes an \mathcal{R}_1 language L , provided $L\tau$ is its set of final states. By Proposition 6.3, all s-variables in the same equivalence class may be replaced by a single variable. Now define a linear system of inequalities \mathfrak{L} as follows:

Definition 6.4. The construction of the linear system of inequalities $\mathfrak{L} = \mathfrak{L}(L)$ for a given \mathcal{R}_1 language L .

- (1) Take the formal expressions $\hat{p}_{\mathbf{v}\theta} = \hat{r}_0 + \hat{r}_{\mathbf{v}[1]} + \hat{r}_{\mathbf{v}[2]} + \dots + \hat{r}_{\mathbf{v}} + \hat{g}_{\mathbf{v}}$ for all $\mathbf{v} \in \mathcal{F}(A)$;
- (2) Obtain linear expressions $\mathfrak{L}(\mathbf{v})$ from $\{\hat{p}_{\mathbf{v}\theta} \mid \mathbf{v} \in \mathcal{F}(A)\}$ in the following way; all s-variables in the same equivalence class $[\hat{r}]$ are replaced by a single variable denoted $\mathfrak{L}(\hat{r})$, while any two s-variables in different equivalence classes are replaced by different variables;

- (3) Introduce yet another variables p_1 and p_2 . For any $\mathbf{v} \in \mathcal{F}(A)$, if $\mathbf{v} \in L\tau$, construct an inequality $\mathfrak{L}(\mathbf{v}) \geq p_2$, otherwise construct an inequality $\mathfrak{L}(\mathbf{v}) \leq p_1$;
- (4) Append the system by an inequality $p_1 < p_2$.

If a MM-BQFA \mathcal{S} recognizes an \mathcal{R}_1 language L , then the linear system of inequalities \mathfrak{L} is consistent. Thus we have established the following result.

Theorem 6.5. *Suppose L is an \mathcal{R}_1 language. If the linear system \mathfrak{L} is not consistent, then L cannot be recognized by any MM-BQFA.*

Therefore, if the linear system \mathfrak{L} is not consistent, then L cannot be recognized by any MM-QFA, DH-PRA or EQFA as well.

Corollary 6.6. $\mathcal{L}(\text{MM-BQFA}) \subsetneq \mathcal{ER}$.

Proof. Consider an \mathcal{R}_1 language $L = \{\mathbf{ab}, \mathbf{bac}\}$ over alphabet $A = \{a, b, c\}$. Among others, the system \mathfrak{L} has the following inequalities:

$$\begin{aligned} \mathfrak{L}(\mathbf{ab}) &= x_0 + x_a + x_{ab} + y_{ab} && \geq p_2 \\ \mathfrak{L}(\mathbf{bac}) &= x_0 + x_b + x_{ba} + x_{abc} + y_{abc} && \geq p_2 \\ \mathfrak{L}(\mathbf{ba}) &= x_0 + x_b + x_{ba} + y_{ab} && \leq p_1 \\ \mathfrak{L}(\mathbf{abc}) &= x_0 + x_a + x_{ab} + x_{abc} + y_{abc} && \leq p_1 \\ &&& p_1 < p_2 \end{aligned}$$

The above inequalities define a system that is not consistent. Hence \mathfrak{L} is not consistent as well. So by Theorem 6.5 L cannot be recognized by any MM-BQFA. Therefore Theorem 5.3 implies that $\mathcal{L}(\text{MM-BQFA}) \subsetneq \mathcal{ER}$. \square

To prove the statement converse to Theorem 6.5, we need to indicate some of the properties of the obtained system \mathfrak{L} . The converse statement itself will be proved in Section 7 (Theorem 7.2).

Consider the inequalities in the system. Let y_A be the unique variable $\mathfrak{L}(\widehat{g}_{\mathbf{w}})$, such that $\mathbf{w}\omega = A$. Except for the inequality $p_1 < p_2$, the left-hand side of any inequality has the form $\mathfrak{L}(\mathbf{v}) = x_0 + x_{\mathbf{v}[1]} + x_{\mathbf{v}[2]} + \dots + x_{\mathbf{v}} + y_{\mathbf{v}}$, where $\mathbf{v} \in \mathcal{F}(A)$, $x_0 = \mathfrak{L}(\widehat{r}_0)$, $x_{\mathbf{v}[j]} = \mathfrak{L}(\widehat{r}_{\mathbf{v}[j]})$, $1 \leq j \leq |\mathbf{v}|$, and $y_{\mathbf{v}} = \mathfrak{L}(\widehat{g}_{\mathbf{v}})$.

The only possible coefficients of variables in any linear inequality are -1 , 0 and 1 . Denote by $Z = \{x_0, z_1, \dots, z_s, y_1, \dots, y_t, p_1, p_2\}$ the set of all the variables in the system \mathfrak{L} , where z_i are variables of the form $x_{\mathbf{v}[j]}$, and y_i are variables of the form $y_{\mathbf{v}}$. Denote by N the total number of variables.

Let $M = |A| + 2$, which is the maximal number of variables (with nonzero coefficients) in any expression $\mathfrak{L}(\mathbf{v})$. Each expression $\mathfrak{L}(\mathbf{v})$ has exactly one variable y_i . If two expressions $\mathfrak{L}(\mathbf{v})$ and $\mathfrak{L}(\mathbf{v}')$ share the same variable y_i , then $\mathbf{v} \sim_{\omega} \mathbf{v}'$, so $\mathfrak{L}(\mathbf{v})$ and $\mathfrak{L}(\mathbf{v}')$ have the same number of variables. So it is possible to denote by $n(y_i)$ the number of variables in any corresponding expression $\mathfrak{L}(\mathbf{v})$. Let $d(y_i) = M - n(y_i) + 1$.

Proposition 6.7. *The system \mathfrak{L} is consistent if and only if it has a solution where all the variables are assigned nonnegative real values.*

Proof. Let c_0, \dots, c_{N-1} be some real numbers. Let C be any real constant. Any inequality in the system can be written in one of the three forms, namely,

$$x_0 + x_{\mathbf{v}[1]} + x_{\mathbf{v}[2]} + \dots + x_{\mathbf{v}} + y_{\mathbf{v}} \geq p_2, \quad (1)$$

$$x_0 + x_{\mathbf{v}[1]} + x_{\mathbf{v}[2]} + \dots + x_{\mathbf{v}} + y_{\mathbf{v}} \leq p_1, \quad (2)$$

$$p_1 < p_2. \quad (3)$$

The inequalities above are satisfied if and only the following inequalities are satisfied;

$$x_0 + x_{\mathbf{v}[1]} + x_{\mathbf{v}[2]} + \dots + x_{\mathbf{v}} + y_{\mathbf{v}} + CM \geq p_2 + CM, \quad (4)$$

$$x_0 + x_{\mathbf{v}[1]} + x_{\mathbf{v}[2]} + \dots + x_{\mathbf{v}} + y_{\mathbf{v}} + CM \leq p_1 + CM, \quad (5)$$

$$p_1 + CM < p_2 + CM. \quad (6)$$

Note that

$$\begin{aligned} x_0 + x_{\mathbf{v}[1]} + x_{\mathbf{v}[2]} + \dots + x_{\mathbf{v}} + y_{\mathbf{v}} + CM &= (x_0 + C) + (x_{\mathbf{v}[1]} + C) + \\ &\quad + (x_{\mathbf{v}[2]} + C) + \dots + (x_{\mathbf{v}} + C) + (y_{\mathbf{v}} + Cd(y_{\mathbf{v}})) \end{aligned}$$

Therefore the system \mathfrak{L} has a solution

$$\begin{aligned} \{x_0 &= c_0, \\ z_1 &= c_1, \dots, z_s = c_s, \\ y_1 &= c_{s+1}, \dots, y_t = c_{N-3}, \\ p_1 &= c_{N-2}, p_2 = c_{N-1}\} \end{aligned} \quad (7)$$

if and only if it has a solution

$$\begin{aligned} \{x_0 &= c_0 + C, \\ z_1 &= c_1 + C, \dots, z_s = c_s + C, \\ y_1 &= c_{s+1} + Cd(y_1), \dots, y_t = c_{N-3} + Cd(y_t), \\ p_1 &= c_{N-2} + CM, p_2 = c_{N-1} + CM\}. \end{aligned} \quad (8)$$

Suppose the system \mathfrak{L} is consistent and has a solution (7). Let $c_{\min} = \min\{0, c_0, \dots, c_{N-1}\}$. Take $C = -c_{\min}$. By construction, C is a nonnegative real number, such that for all i , $c_i + C$ is also nonnegative. Now (8) is the solution of the system \mathfrak{L} such that all the variables are assigned nonnegative values. \square

Proposition 6.8. *The system \mathfrak{L} is consistent if and only if it has a solution where all the variables are assigned nonnegative real values and $x_0 = 0$, $y_A = 0$.*

Proof. Suppose \mathfrak{L} is consistent. By Proposition 6.7, the system has a solution (7), where for all i $c_i \geq 0$. We first prove that there exists a solution where $x_0 = 0$.

If $L \neq A^*$, there exists $\mathbf{v} \in \mathcal{F}(A)$ such that the inequality $\mathfrak{L}(\mathbf{v}) \leq p_1$ is part of the system \mathfrak{L} . Since x_0 is part of $\mathfrak{L}(\mathbf{v})$, $c_0 \leq c_{N-2} < c_{N-1}$. So the system has a solution $\{x_0 = 0, z_1 = c_1, \dots, z_s = c_s, y_1 = c_{s+1}, \dots, y_t = c_{N-3}, p_1 = c'_{N-2}, p_2 =$

$c'_{N-1}\}$, where $c'_{N-2} = c_{N-2} - c_0$ and $c'_{N-1} = c_{N-1} - c_0$. Otherwise, if $L = A^*$, take the solution $\{x_0 = 0, z_1 = 0, \dots, z_s = 0, y_1 = 1, \dots, y_t = 1, p_1 = 0, p_2 = 1\}$.

Next, we prove that there exists a solution where $x_0 = 0, y_A = 0$. The left-hand side of any inequality that contains the variable y_A is of the form $x_0 + x_{\mathbf{v}[1]} + x_{\mathbf{v}[2]} + \dots + x_{\mathbf{v}} + y_A$, where $\mathbf{v}\omega = A$. Any inequality in the system either contains the variable y_A and a single variable $x_{\mathbf{v}}$, such that $\mathbf{v}\omega = A$, or contains none of them. Assume that y_t is the variable y_A and $z_{s-|A|+1}, \dots, z_s$ are all the variables of the form $x_{\mathbf{v}}$, such that $\mathbf{v}\omega = A$. Since the system has a solution $\{x_0 = 0, z_1 = c_1, \dots, z_{s-|A|} = c_{s-|A|}, z_{s-|A|+1} = c_{s-|A|+1}, \dots, z_s = c_s, y_1 = c_{s+1}, \dots, y_{t-1} = c_{N-4}, y_t = c_{N-3}, p_1 = c'_{N-2}, p_2 = c'_{N-1}\}$, it also has a solution where $z_{s-|A|+1} = c_{s-|A|+1} + c_{N-3}, \dots, z_s = c_s + c_{N-3}, y_t = 0$ and other variables keep their previous values. \square

Proposition 6.9. *The system \mathcal{L} is consistent if and only if it has a solution where all the variables are assigned real values from 0 to 1 and $x_0 = 0, y_A = 0$.*

Proof. Suppose \mathcal{L} is consistent. By Proposition 6.8, the system has a solution (7), where for all i $c_i \geq 0$ and $x_0 = 0, y_A = 0$. Assume that y_t is the variable y_A . Let $D = \max\{c_i\}$. Since $p_1 < p_2, D > 0$. So the solution (7) may be divided by D and the system \mathcal{L} has a solution

$$\begin{aligned} \{x_0 = 0, \\ z_1 = c_1/D, \dots, z_s = c_s/D, \\ y_1 = c_{s+1}/D, \dots, y_{t-1} = c_{N-4}/D, y_A = 0 \\ p_1 = c_{N-2}/D, p_2 = c_{N-1}/D\}. \end{aligned} \quad (9)$$

The solution (9) assigns to all the variables real values from 0 to 1. \square

Proposition 6.10. *The system \mathcal{L} is consistent if and only if it has a solution where $x_0 = 0, y_A = 0, 0 \leq p_1, p_2 \leq 1$ and all the other variables $z_1, \dots, z_s, y_1, \dots, y_{t-1}$ are assigned real values from 0 to $1/|A|$.*

Proof. Suppose \mathcal{L} is consistent. By Proposition 6.9, the system has a solution (9), where for all i $0 \leq c_i \leq 1$. For any i , let $c'_i = c_i/D$. Let $c = \max\{c'_1, \dots, c'_{N-4}\}$. If $c'_{N-1} < c|A|$ then the solution is divided by $c|A|$. Otherwise, if $c'_{N-1} \geq c|A|$ then $c \leq 1/|A|$ and no scaling is necessary. \square

7 Construction of DH-PRA for \mathcal{R}_1 languages

In this section, a method will be provided that allows to construct a DH-PRA for any \mathcal{R}_1 language L that generates a consistent system of linear inequalities. Since MM-BQFA is a generalization of DH-PRA, this implies the construction of MM-BQFA as well. Recall σ is a natural morphism from $\mathcal{F}(A)$ to $\mathcal{P}(A)$, defined in subsection 3.3.

Preparation of a linear programming problem. Consider an \mathcal{R}_1 language L over alphabet A . Construct the respective system of linear inequalities \mathcal{L} . Obtain a system \mathcal{L}_1 by supplementing \mathcal{L} with additional inequalities that enforce the

constraints expressed in Proposition 6.10, according to which \mathfrak{L} is consistent if and only if \mathfrak{L}_1 is consistent. Obtain a system \mathfrak{L}'_1 by replacing in \mathfrak{L}_1 the inequality $p_1 < p_2$ by $p_1 \leq p_2$. The linear programming problem, denoted \mathfrak{P} , is to maximize $p_2 - p_1$ according to the constraints expressed by \mathfrak{L}'_1 .

Since \mathfrak{L}'_1 is homogenous, it always has a solution where $p_1 = p_2$. Since the solution polytope of \mathfrak{L}'_1 is bounded, \mathfrak{P} always has an optimal solution. Obviously, if the optimal solution yields $p_1 = p_2$, then \mathfrak{L}_1 is not consistent and therefore, by Theorem 6.5, a DH-PRA that recognizes L does not exist. Otherwise, if the optimal solution yields $p_1 < p_2$, then \mathfrak{L}_1 is consistent.

Automata derived from the free semilattice $\mathcal{P}(A)$. Assume \mathfrak{L}_1 is consistent, so we are able to obtain a solution of \mathfrak{P} where $p_1 < p_2$. Given any expression Z of variables from \mathfrak{L}_1 , let $\mathfrak{P}(Z)$ - the value which is assigned to Z by solving \mathfrak{P} . First, we use the obtained solution to construct probabilistic automata \mathcal{A}_i , $1 \leq i \leq |A|$. Those automata are not probabilistic reversible. Similarly as in the "decide-and-halt" model, the constructed automata have accepting, rejecting and non-halting states. Any input word is appended by the end-marker $\$$. The initial end-marker $\#$ is not used for those automata themselves. Any automaton \mathcal{A}_i is a tuple $(Q_i, A \cup \{\$, \delta_i, s_i, \delta_i)$, where Q_i is a set of states, s_i - an initial state and δ_i - a transition function $Q \times A \times Q \rightarrow [0, 1]$, so $\delta_i(q, a, q')$ is a probability of transit from q to q' on reading input letter a . \mathcal{A}_i is constructed as follows.

- (1) Take the deterministic automaton $(\mathcal{P}(A), A, \emptyset, \cdot)$, remove all the states at level greater or equal to i . The remaining states are defined to be non-halting. The state \emptyset is initial, it is the only state of \mathcal{A}_i at level 0. For any

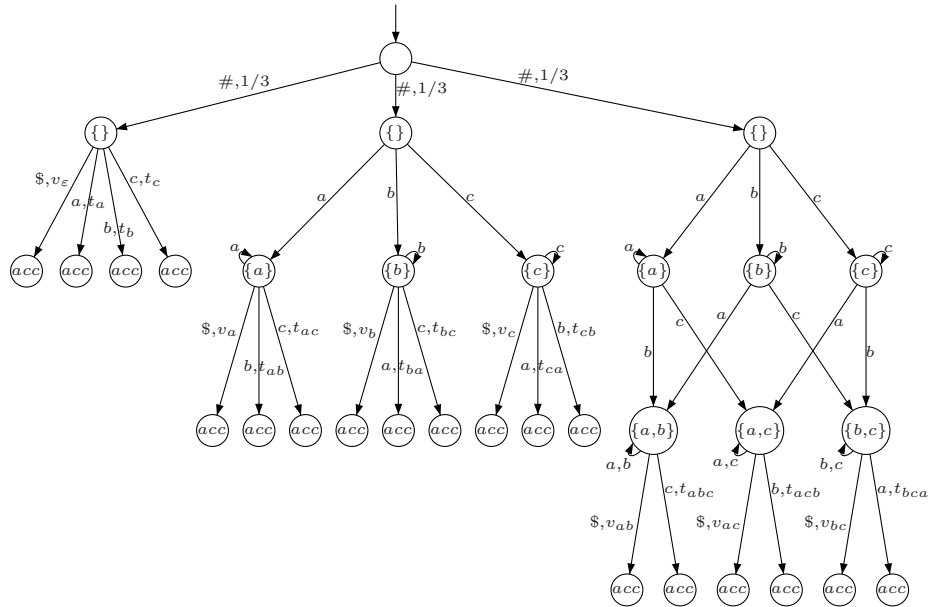


Fig. 3. An automaton \mathcal{A} over alphabet $\{a, b, c\}$, the rejecting states are not shown.

- a in A and state \mathbf{s} at levels $\{0, \dots, i-2\}$, $\delta_i(\mathbf{s}, a, \mathbf{s} \cdot a) = 1$. For any state \mathbf{s} at level $i-1$ and any a in \mathbf{s} , $\delta_i(\mathbf{s}, a, \mathbf{s}) = 1$.
- (2) For any non-halting state \mathbf{s} at levels $\{0, \dots, i-2\}$, add a rejecting state $(\mathbf{s}\$)_{rej}$. Let $\delta_i(\mathbf{s}, \$, (\mathbf{s}\$)_{rej}) = 1$.
 - (3) For any state \mathbf{s} at level $i-1$, add $|A| - |\mathbf{s}| + 1$ accepting states $(\mathbf{s}a)_{acc}$, $a \in (A \setminus \mathbf{s}) \cup \{\$\}$. Also add $|A| - |\mathbf{s}| + 1$ rejecting states $(\mathbf{s}a)_{rej}$, $a \in (A \setminus \mathbf{s}) \cup \{\$\}$.
 - (4) If $a \in A \setminus \mathbf{s}$, any element $\mathbf{s}'a$ in $(\mathbf{s}\sigma^{-1})a$ defines \mathbf{s} -variables in the same equivalence class $[\hat{r}_{\mathbf{s}'a}]$. Let $x_{\mathbf{s}'a} = \mathfrak{L}(\hat{r}_{\mathbf{s}'a})$ and $c_{\mathbf{s}'a} = \mathfrak{B}(x_{\mathbf{s}'a})$. Any element \mathbf{s}' in $\mathbf{s}\sigma^{-1}$ defines \mathbf{s} -variables in the same equivalence class $[\hat{g}_{\mathbf{s}'}]$. Let $y_{\mathbf{s}'} = \mathfrak{L}(\hat{g}_{\mathbf{s}'})$ and $d_{\mathbf{s}'} = \mathfrak{B}(y_{\mathbf{s}'})$.
 - (5) Define missing transitions for the states at level $i-1$. For any state \mathbf{s} at level $i-1$ and any a in $A \setminus \mathbf{s}$, let $t_{\mathbf{s}'a} = \delta_i(\mathbf{s}, a, (\mathbf{s}a)_{acc}) = c_{\mathbf{s}'a}|A|$ and $\delta_i(\mathbf{s}, a, (\mathbf{s}a)_{rej}) = 1 - t_{\mathbf{s}'a}$. Let $v_{\mathbf{s}'} = \delta_i(\mathbf{s}, \$, (\mathbf{s}\$)_{acc}) = d_{\mathbf{s}'}|A|$ and $\delta_i(\mathbf{s}, \$, (\mathbf{s}\$)_{rej}) = 1 - v_{\mathbf{s}'}$.
 - (6) Formally, we would need the transitions outgoing the halting states, those are left undefined.

Consider an automaton \mathcal{A} (Figure 3), which with the same probability $1/|A|$ executes any of the automata $\mathcal{A}_1, \dots, \mathcal{A}_{|A|}$ (i.e., it uses the initial end-marker $\#$ to transit to initial states of any of those automata). By construction of $\mathcal{A}_1, \dots, \mathcal{A}_{|A|}$, the automaton \mathcal{A} accepts any word $\mathbf{u} \in A^*$ with probability $\mathfrak{P}(\mathfrak{L}(\mathbf{u}\tau))$. Since for any word $\mathbf{u} \in L$, $\mathfrak{P}(\mathfrak{L}(\mathbf{u}\tau)) \geq \mathfrak{P}(p_2)$, and for any word $\mathbf{w} \notin L$, $\mathfrak{P}(\mathfrak{L}(\mathbf{w}\tau)) \leq \mathfrak{P}(p_1)$, the automaton \mathcal{A} recognizes the language L .

Construction of a DH-PRA. In order to construct a DH-PRA recognizing L , it remains to demonstrate that any of the automata $\mathcal{A}_1, \dots, \mathcal{A}_{|A|}$ may be simulated by some DH probabilistic reversible automata, that is, for any automaton \mathcal{A}_i , it is possible to construct a sequence of DH-PRA $\mathcal{S}_{i,n}$, where $n \geq 1$, such that $p_{\mathbf{w}, \mathcal{S}_{i,n}}$ converges uniformly to $p_{\mathbf{w}, \mathcal{A}_i}$ on A^* as $n \rightarrow \infty$.

An automaton $\mathcal{A}_i = (Q_i, A \cup \{\$\}, s_i, \delta_i)$ is used to construct a DH-PRA $\mathcal{S}_{i,n} = (Q_{i,n}, A \cup \{\$\}, s_i, \delta_{i,n})$ as described next. Initially $Q_{i,n}$ is empty. Do the following.

- (1) For any non-halting state \mathbf{s} at level j , $0 \leq j \leq i-1$, supplement $\mathcal{S}_{i,n}$ with non-halting states denoted \mathbf{s}_k , where $1 \leq k \leq n^j$.
- (2) For any non-halting state \mathbf{s} at level j , $0 \leq j < i-1$, supplement $\mathcal{S}_{i,n}$ with rejecting states $(\mathbf{s}\$)_{rej,k}$, where $1 \leq k \leq n^j$.
- (3) For any non-halting state \mathbf{s} at level $i-1$, accepting state $(\mathbf{s}a)_{acc}$ and rejecting state $(\mathbf{s}a)_{rej}$, where $a \in (A \setminus \mathbf{s}) \cup \{\$\}$, supplement $\mathcal{S}_{i,n}$ with accepting states $(\mathbf{s}a)_{acc,k}$ and rejecting states $(\mathbf{s}a)_{rej,k}$, where $1 \leq k \leq n^{i-1}$.

It remains to define the transitions. For any non-halting state \mathbf{s} of \mathcal{A}_i at level j , $1 \leq j \leq i-1$, the states in $\{\mathbf{s}_k\}$ are grouped into n^{j-1} disjoint subsets with n states in each, so that any state in $\{\mathbf{s}_k\}$ may be denoted as $\mathbf{s}_{l,m}$, where $1 \leq l \leq n^{j-1}$ and $1 \leq m \leq n$.

For any letter a in A , consider all pairs of non-halting states \mathbf{s}, \mathbf{t} of \mathcal{A}_i such that $\mathbf{s} \neq \mathbf{t}$ and $\delta_i(\mathbf{s}, a, \mathbf{t}) = 1$. For any fixed k and any l and m , $1 \leq l, m \leq n$,

define $\delta_{i,n}(\mathbf{s}_k, a, \mathbf{s}_k) = \delta_{i,n}(\mathbf{s}_k, a, \mathbf{t}_{k,m}) = \delta_{i,n}(\mathbf{t}_{k,m}, a, \mathbf{s}_k) = \delta_{i,n}(\mathbf{t}_{k,l}, a, \mathbf{t}_{k,m}) = 1/(n+1)$.

For any non-halting state \mathbf{s} of \mathcal{A}_i at level j , $0 \leq j < i-1$, $\delta_{i,n}(\mathbf{s}_k, \$, (\mathbf{s}\$)_{rej,k}) = 1$, $\delta_{i,n}((\mathbf{s}\$)_{rej,k}, \$, \mathbf{s}_k) = 1$. For the same $(\mathbf{s}\$)_{rej,k}$ and any other letter b in $A \cup \{\$\}$, define $\delta_{i,n}((\mathbf{s}\$)_{rej,k}, b, (\mathbf{s}\$)_{rej,k}) = 1$.

For any non-halting state \mathbf{s} of \mathcal{A}_i at level $i-1$ and $a \in (A \setminus \mathbf{s}) \cup \{\$\}$, the transitions induced by a among \mathbf{s}_k , $(\mathbf{s}a)_{acc,k}$, $(\mathbf{s}a)_{rej,k}$ are defined by the matrix $\begin{pmatrix} 0 & 0 & 1 \\ r_1 & r_2 & 0 \\ r_2 & r_1 & 0 \end{pmatrix}$, where $r_1 = \delta_i(\mathbf{s}, a, (\mathbf{s}a)_{acc})$, $r_2 = \delta_i(\mathbf{s}, a, (\mathbf{s}a)_{rej})$. The first, second and third rows and columns are indexed by \mathbf{s}_k , $(\mathbf{s}a)_{acc,k}$, $(\mathbf{s}a)_{rej,k}$, respectively. Note that $r_1 + r_2 = 1$. For the same $(\mathbf{s}a)_{acc,k}$, $(\mathbf{s}a)_{rej,k}$ and any other letter b in $A \cup \{\$\}$, define $\delta_{i,n}((\mathbf{s}a)_{acc,k}, b, (\mathbf{s}a)_{acc,k}) = \delta_{i,n}((\mathbf{s}a)_{rej,k}, b, (\mathbf{s}a)_{rej,k}) = 1$.

We have defined all the non-zero transitions for $\mathcal{S}_{i,n}$. By construction, the transition matrices induced by any letter a in $A \cup \{\$\}$ are doubly stochastic.

Lemma 7.1. *For any i , $1 \leq i \leq |A|$, $p_{\mathbf{w}, \mathcal{S}_{i,n}}$ converges uniformly to $p_{\mathbf{w}, \mathcal{A}_i}$ on A^* as $n \rightarrow \infty$.*

Proof. Let $\mathbf{w} \in A^*$ and $p = p_{\mathbf{w}, \mathcal{A}_i}$. Assume $\mathbf{w} = \mathbf{u}\mathbf{y}$, where $|\mathbf{u}\omega| = i-1$ and $|\mathbf{y}| \geq 0$. After reading \mathbf{u} , $\mathcal{S}_{i,n}$ with the same probability $1/(n+1)^{i-1}$ is in one of the $(n+1)^{i-1}$ non-halting states in $\{\mathbf{x}_k \mid 1 \leq k \leq n^{|\mathbf{x}|}, \mathbf{x} \subseteq \mathbf{u}\omega\}$. Among them, there are $\binom{i-1}{l} n^l$ states corresponding to the states of \mathcal{A}_i at level l . So $\mathcal{S}_{i,n}$ has n^{i-1} such states at level $i-1$. Therefore $p_{\mathbf{w}, \mathcal{S}_{i,n}} \geq (\frac{n}{n+1})^{i-1} p$. Also, \mathbf{w} is rejected with probability $q_{\mathbf{w}, \mathcal{S}_{i,n}} \geq (\frac{n}{n+1})^{i-1} (1-p)$. Hence $(\frac{n}{n+1})^{i-1} p \leq p_{\mathbf{w}, \mathcal{S}_{i,n}} \leq 1 - (\frac{n}{n+1})^{i-1} (1-p)$. If $|\mathbf{w}\omega| < i-1$, $p_{\mathbf{w}, \mathcal{S}_{i,n}} = p = 0$.

Since for any j , $0 \leq j \leq |A|$, $\lim_{n \rightarrow \infty} (\frac{n}{n+1})^j = 1$, $p_{\mathbf{w}, \mathcal{S}_{i,n}}$ converges uniformly to $p_{\mathbf{w}, \mathcal{A}_i}$. \square

Now it is possible to construct a DH-PRA $\mathcal{S} = (Q, A \cup \{\#, \$\}, s, \delta)$, which with the same probability $1/|A|$ executes the automata $\mathcal{S}_{1,n}, \dots, \mathcal{S}_{|A|,n}$. The set of states Q is a disjoint union of $Q_1, \dots, Q_{|A|}$. Take the initial state s_i of any $\mathcal{S}_{i,n}$ as the initial state s . For any $a \in A \cup \{\$\}$ and $q_1, q_2 \in Q_i$, $\delta(q_1, a, q_2) = \delta_i(q_1, a, q_2)$. For any initial states s_i and s_j of $\mathcal{S}_{i,n}$ and $\mathcal{S}_{j,n}$, $\delta(s_i, \#, s_j) = 1/|A|$. For any other state q , $\delta(q, \#, q) = 1$. So the transition matrices of \mathcal{S} induced by any letter are doubly stochastic. By Lemma 7.1, \mathcal{S} recognizes L if n is sufficiently large.

Hence we have established the main result of this section:

Theorem 7.2. *Suppose L is an \mathcal{R}_1 language. If the linear system \mathcal{L} is consistent, then L can be recognized by a DH-PRA.*

Therefore, if the linear system \mathcal{L} is consistent, then L can be recognized by a MM-BQFA as well. Moreover, since all of the transition matrices of the constructed DH-PRA are also unitary stochastic, by [15, Theorem 5.2] L can be recognized by an EQFA.

Corollary 7.3. *The class $\mathcal{L}(\text{MM-BQFA})$ is not closed under union and intersection.*

Proof. Consider the \mathcal{R}_1 language $L = \{\mathbf{ab}, \mathbf{bac}\}$ over alphabet $A = \{a, b, c\}$. By Corollary 6.6, L can't be recognized by MM-BQFA.

On the other hand, consider the languages $L_1 = \{\mathbf{ab}\}$ and $L_2 = \{\mathbf{bac}\}$. Systems $\mathfrak{L}(L_1)$ and $\mathfrak{L}(L_2)$ have the same variables as $\mathfrak{L}(L)$. The system $\mathfrak{L}(L_1)$ has a solution where $x_a = 1/2, y_{ab} = 1/2, p_1 = 1/2, p_2 = 1$, and all the other variables equal to 0. The system $\mathfrak{L}(L_2)$ has a solution where $x_b = 1/2, x_{abc} = 1/2, p_1 = 1/2, p_2 = 1$, and all the other variables equal to 0. Therefore by Theorem 7.2 the languages L_1, L_2 are recognized by MM-BQFA. Hence $\mathcal{L}(\text{MM-BQFA})$ is not closed under union. The non-closure under intersection is now implied by closure under complement (Theorem 5.4). \square

8 Construction of MM-QFA for \mathcal{R}_1 languages

The construction of MM-QFA for \mathcal{R}_1 languages has some peculiarities which have to be addressed separately. Specifically, contrary to DH-PRA, EQFA and MM-BQFA, there exist semilattice languages that MM-QFA do not recognize with probability $1 - \epsilon$ [4, Theorem 5] and therefore they can't simulate with the same accepting probabilities the automata $\mathcal{A}_1, \dots, \mathcal{A}_{|A|}$ from Section 7. Nevertheless, MM-QFA still recognize any semilattice language and so a modified construction is still possible.

For any $m \in \mathbb{N}$, let $\alpha(m)$ be the least common multiple of $\{1, 2, \dots, m\}$. Also define $\alpha(0) = 0$. Let O_n - $n \times n$ zero matrix. Let $M_n = (m_{rs}) = (\frac{1}{n})$ and $U_n = (u_{rs}) = \frac{1}{\sqrt{n}}(e^{\frac{2\pi i rs}{n}})$, where $0 \leq r, s \leq n - 1$. M_n is a doubly stochastic matrix and U_n is a unitary matrix that represents the discrete Fourier transform. Obtain an $n \times (n - 1)$ matrix V_n from U_n by removing in U_n its first column. Let V_n^* - the conjugate transpose of V_n . The following lemma will be useful in the construction of MM-QFA.

Lemma 8.1. *The $(2n - 1) \times (2n - 1)$ matrix $H_n = \begin{pmatrix} M_n & V_n \\ V_n^* & O_{n-1} \end{pmatrix}$ is unitary.*

Suppose L is an \mathcal{R}_1 language over alphabet A such that $\mathfrak{L}(L)$ is consistent. As prescribed in Section 7, we construct the automata $\mathcal{A}_1, \dots, \mathcal{A}_{|A|}$, which are the components of the probabilistic automaton \mathcal{A} recognizing L .

Construction of a MM-QFA. For any automaton \mathcal{A}_i , we construct a sequence of MM-QFA $\mathcal{U}_{i,n}$, where $n \geq 1$, such that $n^{\alpha(|A|-1)} p_{\mathbf{w}, \mathcal{U}_{i,n}}$ converges uniformly to $p_{\mathbf{w}, \mathcal{A}_i}$ on A^* as $n \rightarrow \infty$.

An automaton $\mathcal{A}_i = (Q_i, A \cup \{\$\}, s_i, \delta_i)$ is used to construct a MM-QFA $\mathcal{U}_{i,n} = (Q_{i,n}, A \cup \{\$\}, s_i, \delta_{i,n})$ as described next. If $i > 1$, let $c = \frac{\alpha(|A|-1)}{i-1}$, otherwise let $c = 0$. Initially $Q_{i,n}$ is empty. Do the following.

- (1) For any non-halting state \mathbf{s} at level j , $0 \leq j \leq i - 1$, supplement $\mathcal{U}_{i,n}$ with non-halting states \mathbf{s}_k , where $1 \leq k \leq n^{c_j}$. If $|\mathbf{s}| > 0$, new rejecting states \mathbf{s}'_k , $1 \leq k \leq n^{c_j}$, are added to $\mathcal{U}_{i,n}$ as well.

- (2) For any non-halting state \mathbf{s} at level j , $0 \leq j < i - 1$, supplement $\mathcal{U}_{i,n}$ with rejecting states $(\mathbf{s}\$)_{rej,k}$, where $1 \leq k \leq n^{c_j}$.
- (3) For any non-halting state \mathbf{s} at level $i - 1$, accepting state $(\mathbf{s}a)_{acc}$ and rejecting state $(\mathbf{s}a)_{rej}$, where $a \in (A \setminus \mathbf{s}) \cup \{\$\}$, supplement $\mathcal{U}_{i,n}$ with accepting states $(\mathbf{s}a)_{acc,k}$ and rejecting states $(\mathbf{s}a)_{rej,k}$, where $1 \leq k \leq n^{c(i-1)}$.

It remains to define the transitions. For any non-halting state \mathbf{s} of \mathcal{A}_i at level j , $1 \leq j \leq i - 1$, the states in $\{\mathbf{s}_k\}$ are grouped into $n^{c(j-1)}$ disjoint subsets with n^c states in each, so that any state in $\{\mathbf{s}_k\}$ may be denoted as $\mathbf{s}_{l,m}$, where $1 \leq l \leq n^{c(j-1)}$ and $1 \leq m \leq n^c$. The states in $\{\mathbf{s}'_k\}$ are grouped in the same way, so that any state in $\{\mathbf{s}'_k\}$ may be denoted as $\mathbf{s}'_{l,m}$.

For any letter a in A , consider all pairs of non-halting states \mathbf{s}, \mathbf{t} of \mathcal{A}_i such that $\mathbf{s} \neq \mathbf{t}$ and $\delta_i(\mathbf{s}, a, \mathbf{t}) = 1$. For any fixed k and any m , $1 \leq m \leq n^c$, the transitions induced by a among the states $\mathbf{t}_{k,m}, \mathbf{t}'_{k,m}$ and the state \mathbf{s}_k are defined by the matrix H_{n^c+1} ; the first row and column is indexed by \mathbf{s}_k , the next n^c rows and columns by $\mathbf{t}_{k,m}$, and the last n^c rows and columns by $\mathbf{t}'_{k,m}$.

For any non-halting state \mathbf{s} of \mathcal{A}_i at level j , $0 \leq j < i - 1$, $\delta_{i,n}(\mathbf{s}_k, \$, (\mathbf{s}\$)_{rej,k}) = 1$, $\delta_{i,n}((\mathbf{s}\$)_{rej,k}, \$, \mathbf{s}_k) = 1$. For the same $(\mathbf{s}\$)_{rej,k}$ and any other letter b in $A \cup \{\$\}$, define $\delta_{i,n}((\mathbf{s}\$)_{rej,k}, b, (\mathbf{s}\$)_{rej,k}) = 1$.

Consider any non-halting state \mathbf{s} of \mathcal{A}_i at level $i - 1$ and $a \in (A \setminus \mathbf{s}) \cup \{\$\}$. Let $r_1 = \delta_i(\mathbf{s}, a, (\mathbf{s}a)_{acc})$, $r_2 = \delta_i(\mathbf{s}, a, (\mathbf{s}a)_{rej})$. If $i = 1$, let $u_1 = r_1(\frac{1}{n})^{\alpha(|A|-1)}$ and $u_2 = 1 - u_1$. Otherwise, if $i > 1$, let $u_1 = r_1$, $u_2 = r_2$. Note that $u_1 + u_2 = 1$. The transitions induced by a among $\mathbf{s}_k, (\mathbf{s}a)_{acc,k}, (\mathbf{s}a)_{rej,k}$ are defined by the matrix $\begin{pmatrix} 0 & 0 & 1 \\ \sqrt{u_1} & \sqrt{u_2} & 0 \\ \sqrt{u_2} & -\sqrt{u_1} & 0 \end{pmatrix}$. The first, second and third rows and columns are indexed by $\mathbf{s}_k, (\mathbf{s}a)_{acc,k}, (\mathbf{s}a)_{rej,k}$, respectively. For the same $(\mathbf{s}a)_{acc,k}, (\mathbf{s}a)_{rej,k}$ and any other letter b in $A \cup \{\$\}$, define $\delta_{i,n}((\mathbf{s}a)_{acc,k}, b, (\mathbf{s}a)_{acc,k}) = \delta_{i,n}((\mathbf{s}a)_{rej,k}, b, (\mathbf{s}a)_{rej,k}) = 1$.

We have defined all the non-zero transitions for $\mathcal{U}_{i,n}$. By construction, the transition matrices induced by any letter a in $A \cup \{\$\}$ are unitary.

Lemma 8.2. *For any i , $1 \leq i \leq |A|$, $n^{\alpha(|A|-1)}p_{\mathbf{w},\mathcal{U}_{i,n}}$ converges uniformly to $p_{\mathbf{w},\mathcal{A}_i}$ on A^* as $n \rightarrow \infty$.*

Proof. Let $\mathbf{w} \in A^*$ and $p = p_{\mathbf{w},\mathcal{A}_i}$. If $i = 1$, $p_{\mathbf{w},\mathcal{U}_{1,n}} = (\frac{1}{n})^{\alpha(|A|-1)}p$.

Consider the case $i > 1$. Assume $\mathbf{w} = \mathbf{u}\mathbf{y}$, where $|\mathbf{u}\mathbf{y}| = i - 1$ and $|\mathbf{y}| \geq 0$. After reading \mathbf{u} , $\mathcal{U}_{i,n}$ has rejected the input with probability $1 - 1/(n^c + 1)^{i-1}$ and with the amplitude $1/(n^c + 1)^{i-1}$ is in one of the $(n^c + 1)^{i-1}$ non-halting states in $\{\mathbf{x}_k \mid 1 \leq k \leq n^{c|\mathbf{x}|}, \mathbf{x} \subseteq \mathbf{u}\mathbf{y}\}$. Among them, there are $\binom{i-1}{l}n^{cl}$ states corresponding to the states of \mathcal{A}_i at level l . So $\mathcal{U}_{i,n}$ has $n^{c(i-1)}$ such states at level $i - 1$. Therefore $p_{\mathbf{w},\mathcal{U}_{i,n}} \geq (\frac{n^c}{(n^c+1)^2})^{i-1}p$. Also, \mathbf{w} is rejected with probability $q_{\mathbf{w},\mathcal{U}_{i,n}} \geq (\frac{n^c}{(n^c+1)^2})^{i-1}(1 - p) + 1 - (\frac{1}{n^c+1})^{i-1}$. Hence $(\frac{n^c}{(n^c+1)^2})^{i-1}p \leq p_{\mathbf{w},\mathcal{U}_{i,n}} \leq (\frac{n^c}{(n^c+1)^2})^{i-1}p + (\frac{1}{n^c+1})^{i-1} - (\frac{n^c}{(n^c+1)^2})^{i-1}$. Note that $n^{\alpha(|A|-1)} = n^{c(i-1)}$, therefore

$(\frac{n^c}{n^c+1})^{2(i-1)} p \leq n^{\alpha(|A|-1)} p_{\mathbf{w}, \mathcal{U}_{i,n}} \leq (\frac{n^c}{n^c+1})^{2(i-1)} p + (\frac{n^c}{n^c+1})^{i-1} - (\frac{n^c}{n^c+1})^{2(i-1)}$. If $|\mathbf{w}\omega| < i-1$, $p_{\mathbf{w}, \mathcal{U}_{i,n}} = p = 0$.

Since $\lim_{n \rightarrow \infty} (\frac{n^c}{n^c+1})^{i-1} = 1$, $n^{\alpha(|A|-1)} p_{\mathbf{w}, \mathcal{U}_{i,n}}$ converges uniformly to $p_{\mathbf{w}, \mathcal{A}_i}$. \square

Construct a MM-QFA $\mathcal{U}_n = (Q, A \cup \{\#, \$\}, s, \delta)$, which with the same amplitude $1/\sqrt{|A|}$ executes the automata $\mathcal{U}_{1,n}, \dots, \mathcal{U}_{|A|,n}$. The set of states Q is a disjoint union of $Q_1, \dots, Q_{|A|}$. Take the initial state s_i of any $\mathcal{U}_{i,n}$ as the initial state s . For any $a \in A \cup \{\$, \#\}$ and $q_1, q_2 \in Q_i$, $\delta(q_1, a, q_2) = \delta_i(q_1, a, q_2)$. The transitions induced by initial end-marker $\#$ among the initial states s_i of $\mathcal{U}_{i,n}$, $1 \leq i \leq |A|$, are defined by the discrete Fourier transform $U_{|A|}$. For any other state q , $\delta(q, \#, q) = 1$. So the transition matrices of \mathcal{U}_n induced by any letter are unitary.

We are ready to state the main result of the section.

Theorem 8.3. *Suppose L is an \mathcal{R}_1 language. If the linear system \mathfrak{L} is consistent, then L can be recognized by a MM-QFA.*

Proof. If the linear system \mathfrak{L} is consistent, it is possible to construct the corresponding automaton \mathcal{U}_n from above. By Lemma 8.2, $n^{\alpha(|A|-1)} p_{\mathbf{w}, \mathcal{U}_n}$ converges uniformly to $p_{\mathbf{w}, \mathcal{A}}$ on A^* as $n \rightarrow \infty$.

Take $z = \frac{1}{3}(\mathfrak{P}(p_2) - \mathfrak{P}(p_1))$. If n is sufficiently large, for any word $\mathbf{u} \in L$ $n^{\alpha(|A|-1)} p_{\mathbf{u}, \mathcal{U}_n} \geq \mathfrak{P}(p_2) - z$ and for any word $\mathbf{w} \notin L$ $n^{\alpha(|A|-1)} p_{\mathbf{w}, \mathcal{U}_n} \leq \mathfrak{P}(p_1) + z$. Hence for all $\mathbf{u} \in L$ $p_{\mathbf{u}, \mathcal{U}_n} \geq n^{-\alpha(|A|-1)}(\mathfrak{P}(p_2) - z)$ and for all $\mathbf{w} \notin L$ $p_{\mathbf{w}, \mathcal{U}_n} \leq n^{-\alpha(|A|-1)}(\mathfrak{P}(p_1) + z)$. So for any $\mathbf{u} \in L$ and $\mathbf{w} \notin L$ $p_{\mathbf{u}, \mathcal{U}_n} - p_{\mathbf{w}, \mathcal{U}_n} \geq n^{-\alpha(|A|-1)}z$.

Therefore for a sufficiently large fixed n , \mathcal{U}_n recognizes L with bounded error. \square

In summary, we have obtained the following two results:

Theorem 8.4. *Suppose L is an \mathcal{R}_1 language. L can be recognized by MM-QFA if and only if the linear system $\mathfrak{L}(L)$ is consistent.*

Proof. By Theorems 6.5 and 8.3. \square

Theorem 8.5. *MM-QFA, DH-PRA, EQFA and MM-BQFA recognize exactly the same \mathcal{R}_1 languages.*

Proof. By Theorems 6.5, 7.2 and 8.3. \square

9 "Forbidden Constructions"

In [5, Theorem 4.3], Kikusts has proposed "forbidden constructions" for MM-QFA; any regular language whose minimal deterministic finite automaton contains any of these constructions cannot be recognized by MM-QFA. It is actually implied by Theorem 4.7 that the same is true for MM-BQFA and other "decide-and-halt" models from Table 1. Also, by Theorem 5.3 any language that is recognized by a MM-BQFA is contained in \mathcal{ER} . Therefore it is legitimate to ask whether all the \mathcal{ER} languages that do not contain any of the "forbidden constructions" can be recognized by MM-BQFA. In this section, we give a *negative*

answer to this question; we provide an example of an \mathcal{R}_1 language that does not contain any of the "forbidden constructions" and still cannot be recognized by MM-BQFA (and by other "decide-and-halt" models from Table 1).

First, we need a lemma.

Lemma 9.1. *An \mathcal{R}_1 language L has a "forbidden construction" with $n+1$ levels if and only if there exist m, n and words $\mathbf{w}_i, \mathbf{x}_{i,k}, 1 \leq i \leq 2m, 1 \leq k \leq n$, such that*

- (1) $\mathbf{w}_1, \dots, \mathbf{w}_m \in L$;
- (2) $\mathbf{w}_{m+1}, \dots, \mathbf{w}_{2m} \notin L$;
- (3) for any $i, \mathbf{w}_i = \mathbf{x}_{i,1} \dots \mathbf{x}_{i,n}$;
- (4) for any $i, k, \mathbf{x}_{i,k} = \mathbf{x}_{i,k}\tau$;
- (5) for any i, j, k , if $1 \leq k < n$ then $\mathbf{x}_{i,k} \sim_\omega \mathbf{x}_{j,k}$;
- (6) for any $i, \mathbf{w}_i = \mathbf{w}_i\tau$;
- (7) for any k the tuple $(\mathbf{x}_{1,k}, \dots, \mathbf{x}_{m,k})$ is a permutation of $(\mathbf{x}_{m+1,k}, \dots, \mathbf{x}_{2m,k})$.

Proof. Assume an \mathcal{R}_1 language L has a "forbidden construction" of $n+1$ levels. Let l_k - the number of different labels (words) for transitions between levels k and $k+1$. Let $\mathbf{z}_{1,k}, \dots, \mathbf{z}_{l_k,k}$ - the words labeling the transitions from the states at level k to the states at level $k+1$. If $k < n$ there exist transitions labeled $\mathbf{z}'_{1,k}, \dots, \mathbf{z}'_{l_k,k}$ between the states at level k to the states at level $k+1$ such that $\mathbf{z}'_{1,k} \sim_\omega \dots \sim_\omega \mathbf{z}'_{l_k,k}$ and for all $i, \mathbf{z}'_{i,k}\tau = \mathbf{z}'_{i,k}$. If $k = n$ there exist transitions labeled $\mathbf{z}'_{1,n}, \dots, \mathbf{z}'_{l_n,n}$ between the states at level n to the states at level $n+1$ such that for all $i, \mathbf{z}'_{i,n}\tau = \mathbf{z}'_{i,n}$. The states at level $n+1$ are a disjoint union of the sets $D_{1,n}, \dots, D_{l_n,n}$. Therefore the last level has m accepting and m rejecting states, where $m > 0$. Hence there are m words \mathbf{w}'_i in $L, 1 \leq i \leq m$, and m words \mathbf{w}'_j not in $L, m+1 \leq j \leq 2m$. For any $i, 1 \leq i \leq 2m, \mathbf{w}'_i = \mathbf{x}'_{i,1} \dots \mathbf{x}'_{i,n}$, where $\mathbf{x}'_{i,k}$ is equal to some label $\mathbf{z}'_{s,k}$, where $1 \leq s \leq l_k$. So the words \mathbf{w}'_i satisfy the conditions (1-5). Consider the set $D_{s,k}$. Since it has the same number of accepting and rejecting states, $\mathbf{z}'_{s,k}$ occurs the same number of times in the tuples $(\mathbf{x}'_{1,k}, \dots, \mathbf{x}'_{m,k})$ and $(\mathbf{x}'_{m+1,k}, \dots, \mathbf{x}'_{2m,k})$. This implies the condition (7). For all i , let $\mathbf{w}_i = \mathbf{w}'_i\tau$. Since for all i, j and for all k less than $n, \mathbf{x}'_{i,k} \sim_\omega \mathbf{x}'_{j,k}$, the application of τ to the words \mathbf{w}'_i will delete for any k the same letters in $\mathbf{x}'_{1,k}, \dots, \mathbf{x}'_{2m,k}$, thus producing words $\mathbf{x}_{1,k}, \dots, \mathbf{x}_{2m,k}$. So for any $i, \mathbf{w}_i = \mathbf{x}_{i,1} \dots \mathbf{x}_{i,n}$. The words $\mathbf{w}_1, \dots, \mathbf{w}_{2m}$ satisfy all the conditions (1-7).

Now suppose the language L satisfies the conditions (1-7). It is possible to construct a following "forbidden construction". Level 1 consists of a state q_1 and the words $\mathbf{x}_{1,1}, \dots, \mathbf{x}_{2m,1}$. Level 2 consists of states $q_{1,2}, \dots, q_{2m,2}$, such that for any $i, j, q_1\mathbf{x}_{i,1} = q_{i,2}, q_{i,2}\mathbf{x}_{j,1} = q_{i,2}$. Level 2 also has the words $\mathbf{x}_{1,2}, \dots, \mathbf{x}_{2m,2}$. Level $k, 3 \leq k \leq n$, consists of states $q_{1,k}, \dots, q_{2m,k}$, such that for any $i, j, q_{i,k-1}\mathbf{x}_{i,k-1} = q_{i,k}, q_{i,k}\mathbf{x}_{j,k-1} = q_{i,k}$. Level k also has the words $\mathbf{x}_{1,k}, \dots, \mathbf{x}_{2m,k}$. Level $n+1$ consists of states $q_{1,n+1}, \dots, q_{2m,n+1}$, such that for any $i, q_{i,n}\mathbf{x}_{i,n} = q_{i,n+1}$. (Within a "forbidden construction", two states may represent the same state in a minimal deterministic automaton, so it is legible to have the same label in two transitions outgoing a single state.) The states $q_{1,n+1}, \dots, q_{m,n+1}$ are accepting and the states $q_{m+1,n+1}, \dots, q_{2m,n+1}$ are rejecting. \square

Theorem 9.2. *There exists an \mathcal{ER} language that does not contain any of the "forbidden constructions" and still cannot be recognized by MM-BQFA.*

Proof. Consider an \mathcal{R}_1 language $L = \{\mathbf{aedbc}, \mathbf{beca}, \mathbf{beda}, \mathbf{bedac}, \mathbf{each}, \mathbf{eachbd}, \mathbf{eadbc}, \mathbf{ebca}\}$ over alphabet $A = \{a, b, c, d, e\}$. Among others, the system \mathfrak{L} has the following inequalities:

$$\begin{aligned} \mathfrak{L}(\mathbf{aedbc}) &= x_0 + x_a + x_{ae} + x_{aed} + x_{adeb} + x_{abdec} + y_{abcde} \geq p_2 \\ \mathfrak{L}(\mathbf{beda}) &= x_0 + x_b + x_{be} + x_{bed} + x_{bdea} + y_{abde} \geq p_2 \\ \mathfrak{L}(\mathbf{eachbd}) &= x_0 + x_e + x_{ea} + x_{aec} + x_{aceb} + x_{abced} + y_{abcde} \geq p_2 \\ \mathfrak{L}(\mathbf{ebca}) &= x_0 + x_e + x_{eb} + x_{bec} + x_{bcea} + y_{abce} \geq p_2 \\ \mathfrak{L}(\mathbf{aecb}) &= x_0 + x_a + x_{ae} + x_{aec} + x_{aceb} + y_{abce} \leq p_1 \\ \mathfrak{L}(\mathbf{becad}) &= x_0 + x_b + x_{be} + x_{bec} + x_{bcea} + x_{abced} + y_{abcde} \leq p_1 \\ \mathfrak{L}(\mathbf{eadb}) &= x_0 + x_e + x_{ea} + x_{aed} + x_{adeb} + y_{abde} \leq p_1 \\ \mathfrak{L}(\mathbf{ebdac}) &= x_0 + x_e + x_{eb} + x_{bed} + x_{bdea} + x_{abdec} + y_{abcde} \leq p_1 \\ & p_1 < p_2 \end{aligned}$$

$$\begin{array}{lll} \text{Let} & a_1 = x_0 + x_a + x_{ae}, & b_1 = x_{aed} + x_{adeb}, & c_1 = x_{abdec} + y_{abcde}, \\ & a_2 = x_0 + x_b + x_{be}, & b_2 = x_{bed} + x_{bdea}, & c_2 = y_{abde}, \\ & a_3 = x_0 + x_e + x_{ea}, & b_3 = x_{aec} + x_{aceb}, & c_3 = x_{abced} + y_{abcde}, \\ & a_4 = x_0 + x_e + x_{eb}, & b_4 = x_{bec} + x_{bcea}, & c_4 = y_{abce}. \end{array}$$

$$\begin{array}{lll} \text{We obtain inequalities} & a_1 + b_1 + c_1 \geq p_2, & a_1 + b_3 + c_4 \leq p_1, & p_1 < p_2, \\ & a_2 + b_2 + c_2 \geq p_2, & a_2 + b_4 + c_3 \leq p_1, \\ & a_3 + b_3 + c_3 \geq p_2, & a_3 + b_1 + c_2 \leq p_1, \\ & a_4 + b_4 + c_4 \geq p_2, & a_4 + b_2 + c_1 \leq p_1, \end{array}$$

which define a system that is not consistent. Hence \mathfrak{L} is not consistent as well. So by Theorem 6.5 L cannot be recognized by MM-BQFA.

Let us check if L contains any of the "forbidden constructions" from [5, Theorem 4.3]. Since L is \mathcal{R} -trivial idempotent and $|A| = 5$, if L contains some "forbidden construction", by Lemma 9.1, it also must contain a construction with number of levels not larger than 6. Therefore it remains to check the conditions of Lemma 9.1 against constructions with number of levels equal to 3, 4, 5 and 6. In case of 3 levels, it is sufficient to verify that any subset of $\{\mathbf{aedbc}, \mathbf{beca}, \mathbf{beda}, \mathbf{bedac}, \mathbf{each}, \mathbf{eachbd}, \mathbf{eadbc}, \mathbf{ebca}\}$ with at least two elements does not form the words $\mathbf{w}_1, \dots, \mathbf{w}_m$ satisfying all the conditions of Lemma 9.1. Actually, it is sufficient to check only the subsets of $\{\mathbf{aedbc}, \mathbf{each}, \mathbf{eachbd}, \mathbf{eadbc}\}$, $\{\mathbf{beca}, \mathbf{beda}, \mathbf{bedac}, \mathbf{ebca}\}$, $\{\mathbf{aedbc}, \mathbf{beda}, \mathbf{bedac}, \mathbf{eadbc}\}$ and $\{\mathbf{beca}, \mathbf{each}, \mathbf{eachbd}, \mathbf{ebca}\}$. None of these subsets satisfy the conditions of the lemma. The cases with 4, 5 and 6 levels are checked in the same way. So L does not contain any of the "forbidden constructions". \square

References

1. J. Almeida. Finite Semigroups and Universal Algebra. *World Scientific, Singapore*, 1994.

2. A. Ambainis, M. Beaudry, M. Golovkins, A. Ķikusts, M. Mercer, D. Thérien. Algebraic Results on Quantum Automata. *Theory of Computing Systems*, Vol. 39(1), pp. 165-188, 2006.
3. A. Ambainis, R. Freivalds. 1-Way Quantum Finite Automata: Strengths, Weaknesses and Generalizations. *Proc. 39th FOCS*, pp. 332-341, 1998.
4. A. Ambainis, A. Ķikusts. Exact Results for Accepting Probabilities of Quantum Automata. *Theoretical Computer Science*, Vol. 295, pp. 3-25, 2003.
5. A. Ambainis, A. Ķikusts, M. Valdat. On the Class of Languages Recognizable by 1-Way Quantum Finite Automata. *STACS 2001, Lecture Notes in Computer Science*, Vol. 2010, pp. 75-86, 2001.
6. A. Ambainis, J. Watrous. Two-Way Finite Automata with Quantum and Classical States. *Theoretical Computer Science*, Vol. 287, pp. 299-311, 2002.
7. D.S. Bernstein. Matrix Mathematics: Theory, Facts, and Formulas. *Princeton University Press*, 2009.
8. A. Bertoni, C. Mereghetti, B. Palano. Quantum Computing: 1-Way Quantum Automata. *DLT 2003, Lecture Notes in Computer Science*, Vol. 2710, pp. 1-20, 2003.
9. M.P. Bianchi, B. Palano. Behaviours of Unary Quantum Automata. *Fundamenta Informaticae*, Vol. 104, pp. 1-15, 2010.
10. A. Brodsky, N. Pippenger. Characterizations of 1-Way Quantum Finite Automata. *SIAM Journal on Computing*, Vol. 31(5), pp. 1456-1478, 2002.
11. M.D. Choi. Completely Positive Linear Maps on Complex Matrices. *Linear Algebra and Applications*, Vol. 10, pp. 285-290, 1975.
12. M.P. Ciamarra. Quantum Reversibility and a New Model of Quantum Automaton. *FCT 2001, Lecture Notes in Computer Science*, Vol. 2138, pp. 376-379, 2001.
13. I. Dzelme. Galīgs Kvantu Automāts ar Jauktiem Stāvokļiem. *Bakalaura darbs. Latvijas Universitāte, Fizikas un matemātikas fakultāte*, 2003.
14. S. Eilenberg. Automata, Languages and Machines, Vol. B. *Academic Press, New York*, 1976.
15. M. Golovkins, M. Kravtsev. Probabilistic Reversible Automata and Quantum Automata. *COCOON 2002, Lecture Notes in Computer Science*, Vol. 2387, pp. 574-583, 2002.
16. M. Golovkins, M. Kravtsev, V. Kravcevs. On a Class of Languages Recognizable by Probabilistic Reversible Decide-and-Halt Automata. *Theoretical Computer Science*, Vol. 410(20), pp. 1942-1951, 2009.
17. M. Golovkins, J.E. Pin. Varieties Generated by Certain Models of Reversible Finite Automata. *Chicago Journal of Theoretical Computer Science*, Vol. 2010, Article 2, 2010.
18. I. Halperin. The Product of Projection Operators. *Acta Scientiarum Mathematicarum (Szeged)*, Vol. 23(1-2), pp. 96-99, 1962.
19. G.H. Hardy, E.M. Wright. An Introduction to the Theory of Numbers. Fifth Edition. *Oxford University Press*, 1979.
20. J. Hartmanis, R.E. Stearns. Algebraic Structure Theory of Sequential Machines. *Prentice Hall*, 1966.
21. M. Hirvensalo. Quantum Automata with Open Time Evolution. *International Journal of Natural Computing Research*, Vol. 1(1), pp. 70-85, 2010.
22. A. Kondacs, J. Watrous. On The Power of Quantum Finite State Automata. *Proc. 38th FOCS*, pp. 66-75, 1997.
23. G. Kuperberg. The Capacity of Hybrid Quantum Memory. *IEEE Transactions on Information Theory*, Vol. 49-6, pp. 1465 - 1473, 2003.
24. L. Li, D. Qiu, X. Zou, L. Li, L. Wu, P. Mateus. Characterizations of One-Way General Quantum Finite Automata. <http://arxiv.org/abs/0911.3266>, 2010.

25. C. Liu, N. Petulante. On Limiting Distributions of Quantum Markov Chains. <http://arxiv.org/abs/1010.0741>, 2011.
26. M. Mercer. Lower Bounds for Generalized Quantum Finite Automata. *LATA 2008, Lecture Notes in Computer Science, Vol. 5196*, pp. 373-384, 2008.
27. C. Moore, J.P. Crutchfield. Quantum Automata and Quantum Grammars. *Theoretical Computer Science, Vol. 237(1-2)*, pp. 275-306, 2000.
28. A. Nayak. Optimal Lower Bounds for Quantum Automata and Random Access Codes. *Proc. 40th FOCS*, pp. 369-377, 1999.
29. M.A. Nielsen, I.L. Chuang. Quantum Computation and Quantum Information. *Cambridge University Press*, 2000.
30. J. Novotný, G. Alber, I. Jex. Asymptotic Evolution of Random Unitary Operations. *Central European Journal of Physics, Vol. 8(6)*, pp. 1001-1014, 2010.
31. K. Paschen. Quantum Finite Automata using Ancilla Qubits. *Interner Bericht. Universität Karlsruhe, Fakultät für Informatik*, 2000.
32. D. Pérez-García, M.M. Wolf, D. Petz, M.B. Ruskai. Contractivity of Positive and Trace Preserving Maps under L_p Norms. *Journal of Mathematical Physics, Vol. 47-8*, 083506, 2006.
33. J.E. Pin. Varieties of Formal Languages. *North Oxford, London and Plenum, New-York*, 1986.
34. J.E. Pin. $BG = PG$, a Success Story. *NATO Advanced Study Institute. Semigroups, Formal Languages and Groups, J. Fountain (ed.)*, pp. 33-47, *Kluwer Academic Publishers*, 1995.
35. J.E. Pin, H. Straubing, D. Thérien. Small Varieties of Finite Semigroups and Extensions. *J. Austral. Math. Soc. (Series A)*, Vol. 37, pp. 269-281, 1984.
36. M.O. Rabin. Probabilistic Automata. *Information and Control, Vol. 6(3)*, pp. 230-245, 1963.
37. G. Thierrin. Permutation Automata. *Mathematical Systems Theory, Vol. 2(1)*, pp. 83-90, 1968.
38. F. Zhang. Matrix Theory: Basic Results and Techniques. *Springer-Verlag*, 1999.